USAID
FROM THE AMERICAN PEOPLE

All Children Reading–Asia

Analysis of Early Grade Reading Assessment in India

Impact Assessment Report

Submission Date: April 22, 2019

AID-OAA-TO-16-00017 Number: REQ-ASIA-16-00017
TOCOR: Mitch Kirby

Submitted by: **RTI International**
3040 East Cornwallis Road
Research Triangle Park, NC 27709-0155
Tel: (919) 541-6000

Authored by: Maria Dzula, RTI International
3040 East Cornwallis Road
Research Triangle Park, NC 27709-0155
Tel: (919) 541-6000

And

Pratham Education Foundation
ASER Centre
B4/54, Safdarjung Enclave
New Delhi - 110029
Tel: +91 (011) 4602-3612

Table of Contents

Page

List of Figures

List of Tables

List of Acronyms and Abbreviations

ASER	Annual Status of Education Report
CISSD	CARE India: India Solutions for Sustainable Development
CmF	Centre for microFinance
cwpm	correct words per minute
DISE	District Information System for Education
EGR	early grade reading
EGRA	early grade reading assessment
fpc	finite population correction
ORF	oral reading fluency
PPS	probability proportional to size
R2R	Room to Read
RS	random start
SE	standard error
SES	socioeconomic status
SI	sampling interval
SRS	simple random sampling
STIR	Schools and Teachers Innovating for Results
USAID	United States Agency for International Development
USG	United States Government

1 Executive Summary

This report presents the results and findings from the endline data collection and impact assessment of five large education projects within United States (US) Agency for International Development (USAID)/India's portfolio. The purpose of this evaluation activity is to provide data that will allow the Mission to report toward the Global Count and toward the reading indicator *the percent of students who demonstrate reading fluency and comprehension of grade level text at the end of Grade 2 with US Government assistance (USAID Indicator ES 1-1)*.

1.1 Research Background

In September 2017, USAID commissioned RTI and Pratham Education Foundation's (Pratham) Annual Status of Education Report (ASER) Centre to conduct the Analysis of Early Grade Reading Assessment (EGRA) in India activity. Together, RTI and Pratham developed a research plan and modified standard ASER and EGRA instruments to serve the research objective. The five largest education projects from the Mission's portfolio were selected for inclusion into the assessment (***Table 1***). Projects use different approaches and strategies to achieve similar goals – some work through government systems while others are working directly with schools to improve learning outcomes. Project information is summarized in ***Annex B***.

Table 1. Snapshot of the USAID/India-funded early grade reading (EGR) projects included in the evaluation

Project Name	Implementing Partner	Period of Performance
Scaling Up Early Reading Intervention	Room to Read (R2R)	September 2015–September 2020
Nurturing Early Literacy	Centre for microFinance (CmF)	October 2015–September 2019
Teacher Innovations in Practice	Schools and Teachers Innovating for Results (STIR) Education	October 2014–September 2018
Start Early: Read in Time	CARE	July 2014–July 2018
RightToRead	EnglishHelper	September 2015–September 2017; extension in Maharashtra, 2017–2019

1.2 Research Design

The evaluation design to measure gains in reading performance was a difference analysis whereby individual student gains in Standard 2 were calculated (endline score minus initial assessment score). Each project location's impact was evaluated based on the average learning gains achieved in treatment schools as compared with average control school gains. This matched-pairs approach is typically used in longitudinal evaluation designs with a baseline and an endline assessment. Initial data were collected near the beginning of school year between September and October 2017. Data collectors returned to re-test the same students for the endline assessment between the

A key feature of the research design was the collection of longitudinal data. We retested the same students at initial and endline assessment. Impact scores were calculated for each student and averaged by treatment group.

end of February and March 2018 across treatment and control schools.[1]

At initial assessment, the ASER/modified EGRA instrument was administered to a total of 14,370 Standard 2 students randomly selected from 1,191 government primary schools (607 treatment and 584 control) in seven states and 31 districts in India. At endline, 90% of students were tracked and retested, for a total sample of 12,886; with the overall attrition rate at 10%.

Nine project locations were assessed independently. Sample design took into consideration the geographical spread and language of the interventions. Participating schools were randomly selected to create a sample that would be representative of selected project districts. Learners were assessed in one of five languages depending on the language of intervention. Data collectors conducted the field work after attending training and demonstrating mastery of protocol and proper administration of the instruments in schools.

1.3 ASER Reading Assessment Findings

1.3.1 What Percentage of Standard 2 Students are Reading at Each ASER Level at Endline?

The ASER reading assessment indicates the percentage of students who are reading at the beginner level, letter level, word level, and text level.[2] Each student was marked at the highest level at which they could read comfortably.

It is important to note these calculations are straight percentages at endline for each treatment group. Students were not matched to calculate the longitudinal impact.

[I] An important caveat to this impact assessment is that although data may be referred as "baseline" data, they were not collected at the true project baseline. Many projects are well advanced into their second or third year of implementation, with interventions beginning in Standard 1 (the year before the assessment). Therefore, we refer to these data as the initial, or beginning of school year, assessment data to differentiate these data from each project's actual baseline data. Importantly, project schools may have already experienced intervention impact prior to this assessment.

[2] See the Initial Assessment Report for definitions of the levels.

Table 2 shows the percentage of students reading at Standard 2-level text in treatment and control schools for each project location in the endline assessment. The percentages of students at the beginning, letter, word and Standard 1-levels are detailed by project in Section 4. Similar to the initial assessment, results of the ASER reading assessment for endline are wide ranging. The lowest percentage of students reading at Standard 2-level in the treatment group was 1.1% (RightToRead–Maharashtra) with a high of 68% (Nurturing Early Literacy–Maharashtra). Three projects had 30-40% of students reading at the Standard 2-level at endline: Scaling Up Early Learning Intervention–Uttarakhand (31.3%); Start Early: Read in Time–Odisha (31.9%); and Scaling Up Early Learning Intervention–Chhattisgarh (38.8%). The remaining four projects attained percentages of 10% or less of the student sample population.

It is important to note these calculations are straight percentages at endline for each treatment group. Students were not matched to calculate the longitudinal impact.

Table 2. Percent of students reading at Standard 2-level across treatment group by project

Project location	Endline assessment			
	Sample size		Standard 2-level text (% of students)	
	Treatment T	Control C	Treatment T	Control C
Scaling Up Early Learning Intervention–Uttarakhand	881	657	31.3%	14.3%
Scaling Up Early Learning Intervention–Chhattisgarh	835	843	38.8%	18.4%
Nurturing Early Literacy–Rajasthan	579	536	4.0%	6.8%
Nurturing Early Literacy–Karnataka	935	714	5.0%	1.9%
Nurturing Early Literacy–Maharashtra	773	630	68.0%	67.2%
Teacher Innovations in Practice–Uttar Pradesh	795	784	10.0%	10.3%
Start Early: Read in Time–Uttar Pradesh	853	741	9.2%	7.2%
Start Early: Read in Time–Odisha	426	336	31.9%	18.5%
RightToRead–Maharashtra[3]	971	682	1.1%	1.0%

*Standard errors in parentheses

1.3.2 ASER Impact Evaluation Results

Students who were surveyed in the initial assessment were tracked and assessed again during the final assessment (endline).

For each project location, impact on ASER reading assessment is presented by examining the increase in the proportion of students at the Standard 2-level text and the drop in the proportion of students at the Beginner level from initial assessment to final assessment across treatment and control schools. These levels are selected since they are the highest and lowest levels on the ASER reading assessment and projects are expected to reduce the number of students who could not even read letters (are at the Beginner level) and maximize the number of students who could read at the Standard 2 text level.

Table 3 below shows the increase in the proportion of students at the Standard 2 text level as well as the drop in the proportion of students at the Beginner level from initial assessment to final assessment. The difference between treatment and control schools is also shown. To understand whether the difference was significant between treatment and control schools regression analysis was done to capture change in performance at the student level taking advantage of the longitudinal nature of the study.

[3] As noted above, multiple languages were assessed depending on the projects' language of instruction. The RightToRead project works on developing English reading skills. Hence, for this project, the English ASER tool was used. For all other projects, children were assessed in the local language of instruction.

Conclusion. Students in treatment schools performed significantly better than control schools in three project locations:

- Scaling Up Early Learning Intervention in Uttarakhand,
- Scaling Up Early Learning Intervention in Chhattisgarh and,
- Nurturing Early Literacy project in Karnataka.

For the Scaling Up Early Learning Intervention project in Uttarakhand, the difference in the proportion of students who could read the Standard 2-level text from initial to endline assessment in treatment schools (increase of 10.9% points) is significantly higher than this difference in control schools (increase of 3.2% points).

For Scaling Up Early Reading Intervention project in Chhattisgarh, the increase in the proportion of students who could read the Standard 2-level text from initial to endline assessment in treatment schools was 21.2% points. This difference is significantly higher than the difference of 11.1% points in control schools.

For Nurturing Early Literacy project in Karnataka, the difference in the proportion of students who could read the Standard 2-level text from initial assessment to endline in treatment schools (increase of 3.3% points) is significantly higher than this difference in control schools (increase of 0.5% points) confirming the impact of the Nurturing Early Literacy project in Karnataka.

No significant difference was found in the proportion of students reading at beginner level between initial assessment and endline for all project locations.

Table 3. Change in proportion of students at the Standard 2 text and beginner levels from initial assessment to final assessment

ASER subtasks	Treatment group	Average in % points (SE)	IMPACT Difference (T-C)	Effect size
Scaling Up Early Learning Intervention–Uttarakhand				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	10.9 (2.2)	7.7**	0.21
	Control	3.2 (1.9)		
Drop in Beginner level from initial assessment to final assessment	Treatment	6.0 (1.1)	-1.6	-0.06
	Control	7.6 (2.0)		
Scaling Up Early Learning Intervention–Chhattisgarh				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	21.2 (2.3)	10.1***	0.20
	Control	11.1 (1.8)		
Drop in Beginner level from initial assessment to final assessment	Treatment	4.8 (1.2)	-3.6	-0.08
	Control	8.4 (1.9)		
Nurturing Early Literacy–Rajasthan				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	3.7 (0.0)	-1.8	-0.17
	Control	5.5 (1.2)		
Drop in Beginner level from initial assessment to final assessment	Treatment	22.7 (0.4)	6.1	0.05
	Control	16.6 (3.2)		
Nurturing Early Literacy–Karnataka				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	3.3 (0.6)	2.8***	0.15
	Control	0.5 (0.4)		
Drop in Beginner level from initial assessment to final assessment	Treatment	17.1 (1.5)	2.8	0.05
	Control	14.3 (2.4)		
Nurturing Early Literacy–Maharashtra				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	13.3 (2.3)	3.0	0.02
	Control	10.3 (3.8)		
Drop in Beginner level from initial assessment to final assessment	Treatment	0.3 (0.3)	0.1	0.03
	Control	0.3 (0.4)		
Teacher Innovations in Practice–Uttar Pradesh				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	3.7 (1.6)	-2.2	-0.14
	Control	5.9 (1.2)		
Drop in Beginner level from initial assessment to final assessment	Treatment	15.6 (2.8)	-2.4	-0.11
	Control	18.0 (2.9)		

ASER subtasks	Treatment group	Average in % points (SE)	IMPACT Difference (T-C)	Effect size
Start Early Read in Time–Uttar Pradesh				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	4.9 (1.6)	1.4	0.01
	Control	3.4 (1.3)		
Drop in Beginner level from initial assessment to final assessment	Treatment	22.6 (3.5)	6.9	0.03
	Control	15.7 (2.7)		
Start Early Read in Time–Odisha				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	9.7 (2.8)	5.8	0.10
	Control	4.0 (0.0)		
Drop in Beginner level from initial assessment to final assessment	Treatment	10.3 (1.6)	6.1	0.05
	Control	4.3 (0.0)		
RightToRead–Maharashtra (assessed in English)				
Increase in Standard 2-level text from initial assessment to final assessment	Treatment	1.0 (0.0)	-0.1	-0.06
	Control	1.0 (0.8)		
Drop in Beginner level from initial assessment to final assessment	Treatment	16.2 (0.0)	-0.5	-0.04
	Control	16.8 (2.4)		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

1.4 Adapted EGRA Oral Reading Fluency (ORF) Assessment Findings

1.4.1 What Percentage of Standard 2 Students Can Read Fluently With Comprehension as Measured by EGRA ORF?

In addition to the ASER reading assessment students read a second Standard 2-level passage. ***Table 4*** presents the percentage of students reading at or above the reading benchmark at two different time points. Again, it is important to note that these calculations are straight percentages at the initial assessment and endline for each treatment group. Students were not matched to calculate the longitudinal impact.

Nurturing Early Literacy–Maharashtra had the highest percentage of students achieving benchmark at 60%. Scaling Up Early Learning Intervention project locations, Chhattisgarh and Uttarakhand, had over a third of students achieving benchmarks at endline with 37% and 34% respectively. Less than 5% of students are reading at benchmark for three project locations: Nurturing Early Literacy–Rajasthan and Karnataka, and RightToRead.

Table 4. Percent of Students Reaching Benchmark by Project, Treatment Group and Language

Region	Language	Treatment	Percentage of Students Reaching the Benchmark—Initial Assessment	Percentage of Students Reaching the Benchmark—Endline
Scaling Up Early Learning Intervention				
Chhattisgarh	Hindi	Control	4%	11%
		Treatment	12%	37%
Uttarakhand	Hindi	Control	7%	14%
		Treatment	16%	34%
Nurturing Early Literacy				
Rajasthan	Hindi	Control	0%	4%
		Treatment	0%	2%
Maharashtra	Marathi	Control	32%	56%
		Treatment	42%	60%
Karnataka	Kannada	Control	1%	2%
		Treatment	1%	3%
Teacher Innovations in Practice				
Uttar Pradesh	Hindi	Control	3%	7%
		Treatment	4%	8%
		Treatment	4%	8%
Start Early: Read in Time				
Uttar Pradesh	Hindi	Control	2%	4%
		Treatment	3%	6%
Odisha	Odiya	Control	9%	12%
		Treatment	17%	26%
RightToRead				
Maharashtra	English	Control	0%	0%
		Treatment	0%	3%

1.4.2 Adapted EGRA Impact Evaluation Results

Impact on the EGRA is reported based on the mean gains on two key indicators—ORF and reading comprehension—for the treatment vs. control groups. The longitudinal survey design, namely tracking and testing the same student at initial and endline assessments, afforded the evaluation a unique opportunity to measure individual student gains relative to their initial scores. For our longitudinal data, rather than simply averaging scores for the treatment and control groups at initial assessment and subtracting the averaged endline scores, we can see how individual students performed. The difference, typically a gain, in student scores was calculated for each individual student. The individual student gains were then averaged for the treatment group to calculate the mean gain in student scores for treatment. This same procedure was repeated for students in the control group. The difference between treatment and control is the intervention impact, which was calculated as the difference in mean gains between treatment and control. This gain difference represents impact and is shown in ***Table 5***.

The results show that treatment outperformed control in four project locations:

- Scaling Up Early Reading Intervention—Chhattisgarh,
- Scaling Up Early Reading Intervention—Uttarakhand,
- Start Early Read in Time—Odisha, and
- RightToRead—Maharashtra.

The gain differences between treatment and control were significant at these project locations at the 0.01 level. The Scaling Up Early Reading Intervention Project—Chhattisgarh treatment group's mean ORF improved by 7.6 correct words per minute (cwpm) over the control, and the treatment group gained 10.2% in reading comprehension (% correct) over the control. The Scaling Up Early Reading Intervention—Uttarakhand treatment group's mean ORF improved by 5.4 cwpm over the control, and the treatment group also gained 7.7% in reading comprehension (% correct) over the control.

Conclusion. To a large extent, the results of the EGRA impact analysis mimic the ASER findings. Treatment outperformed control in both Scaling Up Early Reading Intervention project locations, and the difference in gains was found to be significant. Similar to ASER, we also measured a small to medium effect size for these project locations.[4]

Unlike ASER, the mean gains for Start Early Read in Time—Odisha were significant at the 0.01 level for the mean gain in ORF (gain difference of 3.9 cwpm for treatment). We also found a significant gain difference in ORF for the treatment group of the RightToRead Maharashtra project. On average, students in the treatment group gained 2.4 cwpm over those in the control group.

Lastly, we found no significant difference in gains between treatment and control for Nurturing Early Literacy—Karnataka.

[4] At the initial assessment, we did not achieve balance between the treatment and control groups, with a larger proportion of students scoring zero in the control group at both project locations.

Table 5. Average student gains in ORF and reading comprehension scores

Subtasks	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Scaling Up Early Reading Intervention Project–Chhattisgarh				
ORF (cwpm)	Control	6.06 (0.6)	7.57**	0.43 Medium
	Treatment	13.63 (0.7)		
Reading comprehension (% correct)	Control	4.93% (0.8%)	10.2%**	0.33 Small
	Treatment	15.12% (1.1%)		
Scaling Up Early Reading Intervention Project–Uttarakhand				
ORF (cwpm)	Control	6.05 (0.7)	5.44**	0.24 Small
	Treatment	11.49 (0.7)		
Reading comprehension (% correct)	Control	5.79% (1%)	7.71%**	0.20 Small
	Treatment	13.5% (1.2%)		
Nurturing Early Literacy–Maharashtra				
ORF (cwpm)	Control	13.9 (0.8)	-1.99*	-0.04
	Treatment	11.9 (0.5)		
Reading comprehension (% correct)	Control	21.09% (1%)	-1.97%	-0.02
	Treatment	19.12% (0.8%)		
Nurturing Early Literacy–Rajasthan				
ORF (cwpm)	Control	5.1 (0.8)	-1.13	-0.02
	Treatment	4 (0)		
Reading comprehension (% correct)	Control	3.95% (0.9%)	-0.84%	-0.01
	Treatment	3.11% (0%)		
Nurturing Early Literacy–Karnataka				
ORF (cwpm)	Control	2.1 (0.4)	1	0.04
	Treatment	3.1 (0.3)		
Reading comprehension (% correct)	Control	2.26% (0.6%)	0.49%	0.01
	Treatment	2.75% (0.4%)		
Teacher Innovations in Practice–Uttar Pradesh				
ORF (cwpm)	Control	3.46 (0.7)	-0.22	-0.04
	Treatment	3.25 (0.6)		
Reading comprehension (% correct)	Control	3.37% (0.8%)	-0.36%	-0.04
	Treatment	3.02% (0.8%)		
Start Early: Read in Time–Odisha				
ORF (cwpm)	Control	3.4 (0.7)	3.94**	0.14
	Treatment	7.3 (0.9)		
Reading comprehension (% correct)	Control	4.25% (1.4%)	2.95%	0.06
	Treatment	7.2% (1.1%)		
Start Early: Read in Time–Uttar Pradesh				
ORF (cwpm)	Control	2.79 (0.8)	0.93	0.06
	Treatment	3.73 (0.7)		

Subtasks	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Reading comprehension (% correct)	Control	1.23% (1%)	2.07%	0.09
	Treatment	3.3% (0.8%)		
RightToRead–Maharashtra				
ORF (cwpm) in English	Control	0.89 (0.4)	2.39**	0.11
	Treatment	3.29 (0.5)		
Reading comprehension (% correct)	Control	1.8% (0.9%)	-0.41%	-0.03
	Treatment	1.38% (0.4%)		
	Treatment	1.38% (0.4%)		

*Significant at the 0.5 level
** Significant at the 0.01 level

1.4.3 Gains in Reading Benchmarks

Treatment outperformed control, and gain differences were significant at the 0.01 level in four project locations:

1. Scaling Up Early Reading Intervention Project—Chhattisgarh,
2. Scaling Up Early Reading Intervention Project—Uttarakhand,
3. Start Early: Read in Time—Odisha, and
4. EnglishHelper—Maharashtra.

The largest gain in the percentage of students reaching benchmark was found for the Chhattisgarh project location, with a gain difference of 18.1%, followed by Uttarakhand with 10.7% (***Table 6***). For Chhattisgarh, 25% more students achieved the reading benchmark in the treatment group, and 6.9% more students reached the benchmark for the control group. The net gain of 18.1% is the intervention impact. However, as mentioned previously, for this project location, a medium effect size was detected between treatment and control at the initial time of assessment. A medium effect size (i.e., one that is greater than 0.2) indicates that a difference between control and treatment groups does not satisfy equivalence.[5] Furthermore, it is not possible to determine whether this difference at the initial assessment is attributable to demographic differences between the control and treatment groups or to gains in the treatment group resulting from the intervention in Standard 1.

Table 6. Percentage of students reaching the reading benchmark

	Treatment	Change in Percentage	Difference (T-C)	Effect Size
Scaling Up Early Reading Intervention Project—Chhattisgarh				
Hindi 35-cwpm benchmark	Control	6.9%	18.1%**	0.30 Medium
	Treatment	25.0%		
Scaling Up Early Reading Intervention Project—Uttarakhand				
Hindi 35-cwpm benchmark	Control	7.0%	10.7%**	0.15
	Treatment	17.7%		
Start Early: Read in Time—Odisha				
Oriya	Control	3.0%	6.7%**	0.10

[5] (Institute of Educational Sciences, U.S. Department of Education, 2014).

	Treatment	Change in Percentage	Difference (T-C)	Effect Size
30-cwpm benchmark	Treatment	9.7%		
RightToRead—Maharashtra				
English 30-cwpm benchmark	Control	0.3%	2.5%**	0.07
	Treatment	2.8%		

**Significant at the 0.01 level

Reading benchmarks for all project locations are presented in ***Section 5***.

The project research design and methodology for measuring impact are detailed in the next section, ***Section 2***. Information on the endline data collection follows in ***Section 3***, with endline results and impact discussed in ***Section 4***. ***Section 5*** provides data on the gains in reading benchmark attainment and is followed by a short conclusion in ***Section 6***.

2 Measuring Impact: Project Research Design

2.1 Research Questions

The central research questions are as follows:

1) What percentage of Standard 2 students are reading at each level (i.e., beginner, letter, word, Standard 1-level text, and Standard 2-level text) for each project at initial and final assessment, as measured by the ASER reading assessment?
2) What percentage of Standard 2 students can read fluently with comprehension for each project at initial and final assessment, as measured by the adapted EGRA ORF subtask?

Final ASER reading results from each project will be provided to USAID to be used for reporting progress toward USAID's Goal 1 target of improved reading for 100 million children. Student results at the final assessment for the ORF and reading comprehension subtasks will be used to set benchmarks for each language against which Indicator ES 1-1 will be calculated.

2.2 Research Design

Of the 13 USAID/India-supported EGR projects, 5 projects are included in this evaluation report (***Table 7***). The evaluation design aimed to measure increases in student reading performance over the course of one school year in treatment and control schools, as measured at the beginning and end of the 2017/2018 school year.

This study was longitudinal at the student level, meaning that the same students were tested twice: once at the beginning of the school year and again at the end of the school year. The initial assessment gathered data from Standard 2 students who had been (or will have been by the final assessment) exposed to a USAID education project (treatment) and Standard 2 students who had not been exposed to any USAID-funded education project (control). The longitudinal design of this study warranted a difference analysis, whereby student reading skill gains were calculated at the two time points. This analysis allows us to generalize the results for the Standard 2 population per project location for both the initial and final data collections and evaluate each project location's impact based on the gains achieved in treatment schools compared with control schools.

As summarized in ***Table 7***, seven states were included in the evaluation: Uttarakhand, Chhattisgarh, Rajasthan, Uttar Pradesh, Odisha, Karnataka, and Maharashtra. In total, nine project locations were evaluated. Sampling was designed to obtain representative estimates for the reading performance of Standard 2 students in each of these project locations. See ***Section 2.3*** for details on sampling.

Students were assessed using the ASER reading test and EGRA's ORF and comprehension subtasks. Each student was assessed orally in the language of instruction being used in the respective project's intervention, but the language of instruction varies across projects (***Table 7***). Instruments were developed and administered in the following languages: 1) Hindi, 2) Marathi, 3) Oriya, 4) Kannada, and 5) English. It is important to note that results across languages cannot be compared given the differences in language complexity and orthography.

Table 7. Summary of the evaluated projects' geographies and languages of assessments

Project Name	Implementer	Geography (State)	Districts Included in Assessment (number in parenthesis)	Language of Assessment
Scaling Up Early Reading Intervention	R2R	Uttarakhand	Almora, Champawat, Dehradun, Udham Singh Nagar (4)	Hindi
		Chhattisgarh	Baloda Bazar, Raipur (2)	
Nurturing Early Literacy	CmF	Rajasthan	Sirohi (1)	Hindi
		Karnataka	Yadgir (1)	Kannada
		Maharashtra	Satara (1)	Marathi
Teacher Innovations in Practice	STIR Education	Uttar Pradesh	Barabanki, Chandauli, Faizabad, Jaunpur, Kanpur City, Lucknow, Mirzapur, Rae Bareilly, Unnao, Varanasi (10)	Hindi
Start Early: Read in Time	CARE	Uttar Pradesh	Bahraich, Balrampur, Gonda, Hardoi, Shravasti (5)	Hindi
		Odisha	Mayurbhanj (1)	Oriya
RightToRead	EnglishHelper	Maharashtra	Nagpur, Latur, Solapur, Pune, Osmanabad, Jalgaon (6)	English

2.3 Sampling Design

As previously mentioned, the five USAID/India-funded EGR projects included in this evaluation are spread across seven states in India. Because of language and other cultural, socioeconomic, and policy-level heterogeneity, it was important to design the sample to produce estimates for each project at the state level. As a result, data were collected across nine project locations spanning seven states and 31 districts. The sampling design provides representative estimates for reading performance of Standard 2 students in each of these project locations. For a detailed discussion on sampling, see ***Annex A***.

For the initial assessment, a total of 14,370 students from Standard 2 were assessed from 1,191 schools (607 treatment and 584 control) across all nine project locations (see the Initial Assessment Report for details). For the endline assessment, the same schools were revisited, and the same students were tracked and assessed again. ***Section 3*** explains the

data collection process for the endline assessment. Approximately 90% of students were tracked and assessed during both visits (10% attrition rate). At endline, 12,886 Standard 2 students (6,402 girls) were retested.

The final sample of students tracked and assessed in the endline assessment and the sample of students assessed across both initial and endline visits by project and state are presented in ***Table 8.***

Table 8. School and student sample by project location

Program	Number of Schools*			Initial Assessment		Endline Assessment						Assessed in both Initial and Endline Assessments					
				Student Sample in Treatment Schools	Student Sample in Control Schools	Student Sample in Treatment Schools			Student Sample in Control Schools			Student Sample in Treatment Schools			Student Sample in Control Schools		
	T	C	Total	All	All	Boys	Girls	All**	Boys	Girls	All**	Boys	Girls	All**	Boys	Girls	All**
Scaling Up Early Reading Intervention—Uttarakhand	90	90	180	974	707	432	449	881	309	347	657	427	450	881	310	342	657
Scaling Up Early Reading Intervention—Chhattisgarh	60	60	120	932	950	434	399	835	399	444	843	432	397	834	395	440	838
Nurturing Early Literacy—Rajasthan	60	60	120	666	591	288	291	579	266	270	536	283	290	573	267	265	532
Nurturing Early Literacy—Karnataka	60	60	120	1,039	783	456	479	935	358	356	714	458	475	933	354	360	714
Nurturing Early Literacy—Maharashtra	70	70	140	814	656	392	381	773	324	306	630	385	386	771	327	301	628
Teacher Innovations in Practice—Uttar Pradesh	70	70	140	896	869	406	384	795	375	408	784	390	383	782	368	390	764
Start Early: Read in Time—Uttar Pradesh	70	70	140	946	826	457	395	853	381	360	741	454	382	848	374	351	729
Start Early: Read in Time—Odisha	60	60	120	497	408	204	209	426	158	166	336	207	211	426	150	169	326
RightToRead—Maharashtra	67	44	111	1,064	752	509	462	971	339	343	682	501	467	968	339	343	682
Total	607	584	1,191	7,828	6,542	3,578	3,449	7,048	2,909	3,000	5,923	3,537	3,441	7,016	2,884	2,961	5,870

* See Annex A for information about the sampling of additional schools.
** The sum of boys and girls does not always add up to “All” because gender was not recorded for some students.

2.4 Research Limitations

This section details a number of important interpretations and limitations to consider when reviewing these impact data, especially as the data are not collected prior to the start of project implementation.

At the initial assessment, it was not possible to conduct the baseline balance assessment. The extent to which the difference between the control and treatment gains can be attributable to the intervention remains unknown.

It is important to note that this assessment was conducted with Standard 2 students; almost all students in the sampled treatment schools participated in intervention programs beginning in Standard 1. Consequently, the gains in reading outcomes described in ***Section 4*** are a combination of students' achievement through the school year and their receipt of the intervention in Standard 1. Therefore, assessing balance for the initial assessment was not possible. The sample design attempted to match up schools between control and treatment blocks; however, treatment schools may already be exhibiting gains in reading fluency and comprehension because of the Standard 1 interventions. Unfortunately, these gains are masked and cannot be attributed to the intervention's impact.

Additional limitations are as follows:

- Initial data collection did not begin until mid-September 2017 and was only completed in mid-October 2017 for some schools. Because endline data collection occurred between February and March 2018, the evaluation measured impact over a 5- to 6-month period. Although the school year was shortened, ASER's previous research showed that September–March is the most productive time of the school year, with the largest learning gains being observed in this period.
- Ideally, control schools should be matched to the learning levels in treatment schools. However, no such secondary data are available at the student or school level. Schools were matched based on the school-level information available from the District Information System for Education (DISE). Because the initial assessment was done after 1 year of intervention (Standard 1 year), we were unable to assess the balance between control and treatment schools using student outcomes.
- USAID Indicator ES 1-1 is typically calculated by conducting cross-sectional analysis of just the treatment group over 1 year. By making the same calculation within a school year, we risked entangling the learning gains attributed to the intervention with any gains typically seen from being in school.
- The same students were retested using the same instrument; therefore, they were familiar with the structure and content of the assessment when being tested at the final assessment. However, if a student's reading abilities had improved by the final assessment, that student might have been able to read further in the passage (i.e., reading new content not read at the initial assessment) and would have received additional questions.
- Spillover effects may be seen. Control schools were sampled from the same blocks (district sub-divisions) as treatment schools to ensure closer matching. Teachers and cluster- and block-level officers may have met and interacted with each other about interventions at block-level meetings, leading to spillover effects between treatment and control schools.
- We may see the Hawthorne effect. Intervention organizations were designated by USAID/India to arrange for permissions for data collection in schools; therefore, sampled schools were notified of their inclusion in the study and provided with the exact days that data collectors would visit the schools to assess students. These notifications might have resulted in the Hawthorne effect, meaning that normal school

practices may have been modified as a reactive measure to knowing that data collectors were visiting the school and conducting student assessments.

- Outliers in the data could be found. While outliers are perfectly acceptable and to be expected, one school in the initial assessment was many standard deviations outside what we would expect. An investigation by ASER revealed that this school was unsuitable for the sample and was removed from initial and endline analysis.
- Lastly, it is important to note that control schools were not devoid of any school intervention except for the intervention in question, as some schools may have had influence from other interventions not part of this research study, or sustained inputs from government.

3 Endline Data Collection

Student performance in reading and comprehension was assessed using the ASER reading assessment and the EGRA ORF subtask developed for India. The same instruments were used for initial and endline assessments, and students were evaluated using the same reading test.[6]

The ASER reading assessment categorizes students in one of five levels: non-reader (beginner), letter level, word level, Standard 1 level, and Standard 2 level. Students were marked at the highest level at which they could read comfortably. The data collector began each assessment on Standard 1-level text, and depending on how the child performed, the child was asked to read the more challenging Standard 2-level text or the less challenging words and, then, the letters subtasks.

The EGRA measures basic skills that a child must possess to eventually be able to read fluently and with comprehension—the ultimate goal of reading, where reading with comprehension is defined as achieving 80% correct on the reading comprehension subtask. For this evaluation, the EGRA portion only included ORF and reading comprehension subtasks. The additional passage was timed using a stopwatch, and students were asked five questions based on the passage.

The initial assessment data provided beginning-of-grade measures of performance for students in Standard 2. Prior to the endline assessment, data collectors received a refresher training using a two-tier, cascade training model. Importantly the training focused on the process of tracking and retesting the same students who were surveyed and assessed in the initial assessment.

A master training was held centrally on February 6–8, 2018, in Jaipur, Rajasthan. During this training, 31 master trainers were trained on the administration of the assessment and the process to track students assessed in the initial assessment. In the following weeks, these master trainers were deployed for the state training, where they trained the actual surveyors (data collectors). Each field team was made up of the following people: data collector, monitor, and supervisor/ASER state team. Details on the training of data collectors and data collection dates are provided in ***Annex D***. Information on quality control measures is included in ***Annex E***.

The endline assessments were carried out between mid-February and end-March 2018, providing end-of-grade points of comparison. As mentioned above, this schedule meant that the reading interventions were implemented for approximately 5–6 months prior to the endline assessment. In addition to the student assessments, the endline data collection also

[6] Detailed descriptions of the reading assessments are available in the Initial Assessment Report. The instruments used for this evaluation are available upon request.

surveyed students with a short student questionnaire regarding their families and home situations. In that section and throughout this report, all summaries use weighted data and, thus, are representative of the entire population of schools from which the samples were drawn. Information on weighting of estimates is included in ***Annex A***.

Students who were surveyed in the initial assessment were tracked and assessed at endline. Tracking sheets with pre-filled information from the initial assessment were created to locate the same students from the initial assessment. For each school in every project location, a separate tracking sheet was generated with Unique Student IDs. These tracking sheets had basic location identifiers for the schools and students, which helped data collectors to find the students in the schools.

For the final assessment, schools were informed about the data collection dates at least one to two days in advance and were requested to ensure students attend the school on the day of assessment as far as possible. For students who were not attending school on the day of final assessment, data collectors attempted to track and assess them in their households. As a result, 90% of students were assessed in both initial and final visit across all project locations. The overall attrition for treatment schools and control schools were very similar at 10.4% and 10.3%, respectively. See ***Annex C*** for attrition for each project. Despite attempts to locate students, 756 students in treatment schools and 637 students in control schools were not tracked. The most widespread reason for students not being tracked was that they were outside of the village or moved away.

4 Impact Results

This section presents the main impact findings from the endline assessment. For each project location, performance on the ASER reading assessment is first presented only for the endline assessment. Next, the movement of students across different reading levels from the initial assessment to endline assessment is examined. By understanding the increase in the proportion of students at the Standard 2-level text and the drop in the proportion of students at the beginner level from initial assessment to final assessment across treatment and control schools we can assess student performance. These levels are selected since they are the highest and lowest levels on the ASER reading assessment. Finally, the performance of students on the ORF and comprehension subtasks is presented for each location.

The results show the proportion of students at each level of the ASER reading assessment at endline and the gains in mean scores on ORF and reading comprehension by control and treatment. The effect size—Cohen's *d*—is provided and was calculated as follows: the mean treatment score minus the mean control score, divided by the pooled standard deviation. The effect size is used to express the magnitude of the intervention impact. Cohen attributed small, medium, and large effect sizes to values of 0.2, 0.5, and 0.8, respectively (Cohen, 1992).

A number of limitations are outlined in ***Section 2.4*** and are important considerations for interpreting impact results. Primarily, what we call the initial, or baseline, assessment was not a true baseline as commonly understood for evaluations. The baseline for the purposes of our research questions was not conducted prior to the start of school interventions, which has major implications for the findings presented below. As outlined in the Initial Assessment Report, it was not possible to ensure balance between control and treatment groups. The Standard 2 students assessed in treatment schools were exposed to intervention efforts starting in Standard 1. Therefore, the findings presented for treatment groups are possibly a result of students' exposure to intervention efforts and, therefore, already show some treatment effect. However, this is only the case if the control and treatment groups were, in

fact, statistically similar before the intervention. Again, because this assessment is taking place long after project interventions began, comparability at baseline cannot be proven. While the sample design attempted to match control schools, socioeconomic (SES) data collected at the initial assessment provided evidence that balance was not achieved across all project locations.

4.1 Scaling Up Early Reading Intervention Project

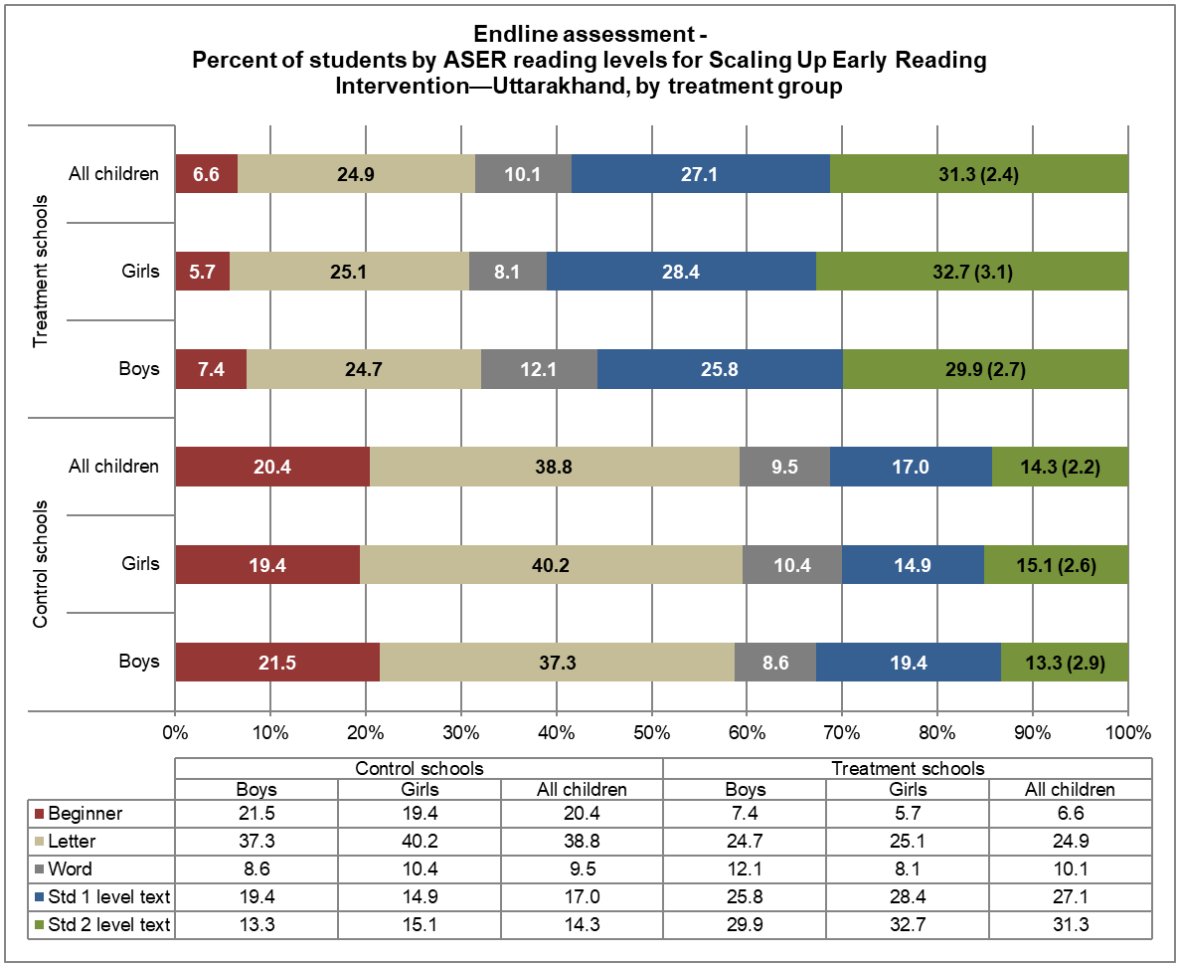
R2R is implementing the Scaling Up Early Reading Intervention project in two states in India: Uttarakhand and Chhattisgarh. Beginning in September 2015, the project is in the third year of a 5-year implementation (through September 2020). The target is to reach 460,000 children in Standards 1–5. For the assessment, students were assessed in Hindi in four districts in Uttarakhand and two districts in Chhattisgarh.

4.1.1 Scaling Up Early Reading Intervention Project—Uttarakhand

(1) ASER Results

Figure 1 shows the performance of Standard 2 students in the endline assessment in treatment and control schools for the Scaling Up Early Reading Intervention project in Uttarakhand. Of students in treatment schools, 31.3% could read the Standard 2-level text, and another 27.1% of students could read the Standard 1-level text but not the Standard 2-level text. In control schools, these proportions are 14.3% and 17%, respectively. In the endline assessment, 20.4% of Standard 2 students in control schools could not even read letters, compared to 6.6% of students in treatment schools.

Figure 1. Percentage of students by ASER reading level at endline (Scaling Up Early Reading Intervention—Uttarakhand)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	21.5	19.4	20.4	7.4	5.7	6.6
Letter	37.3	40.2	38.8	24.7	25.1	24.9
Word	8.6	10.4	9.5	12.1	8.1	10.1
Std 1 level text	19.4	14.9	17.0	25.8	28.4	27.1
Std 2 level text	13.3	15.1	14.3	29.9	32.7	31.3

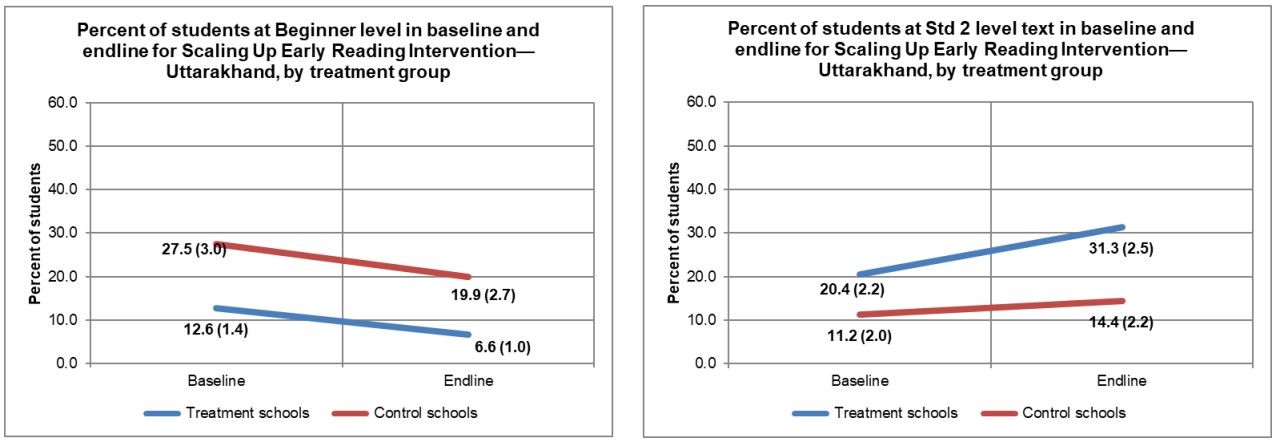
*Standard errors (SEs) in parentheses

To understand the effect of the Scaling Up Early Reading Intervention project in Uttarakhand on students' reading performance, we examined the change in the proportion of students at beginner and Standard 2 text levels from the initial assessment to the endline assessment for treatment and control schools. To understand whether the difference was significant between treatment and control schools regression analysis was done to capture change in performance at the student level taking advantage of the longitudinal nature of the study.

From ***Figure 2***, the proportion of students who could read the Standard 2-level text increased from 20.4% in the initial assessment to 31.3% in the final assessment for treatment schools. The corresponding increase in control schools was from 11.2% to 14.4%. Regression analysis (see Annex G) shows that the difference in the proportion of students who could read the Standard 2-level text from initial to final assessment in treatment schools (10.9% points) is significantly higher than this difference in control schools (3.2% points) confirming the impact of the Scaling Up Early Reading Intervention project in Uttarakhand.

For the Scaling Up Early Reading Intervention project in Uttarakhand, the proportion of students at the Beginner level reduced from 12.6% in the initial assessment to 6.6% in the final assessment for treatment schools. For control schools, the proportion of students who could not even read letters reduced from 27.5% to 19.9%. However, regression analysis (see Annex H) confirms that the difference in the proportion of students who were at the Beginner level from initial to final assessment in treatment schools (6.0% points) is not significantly different from this difference in control schools (7.6% points).

Figure 2. Percentage of students at the beginner and Standard 2 levels at the initial and endline assessments (scaling Up Early Reading Intervention—Uttarakhand)



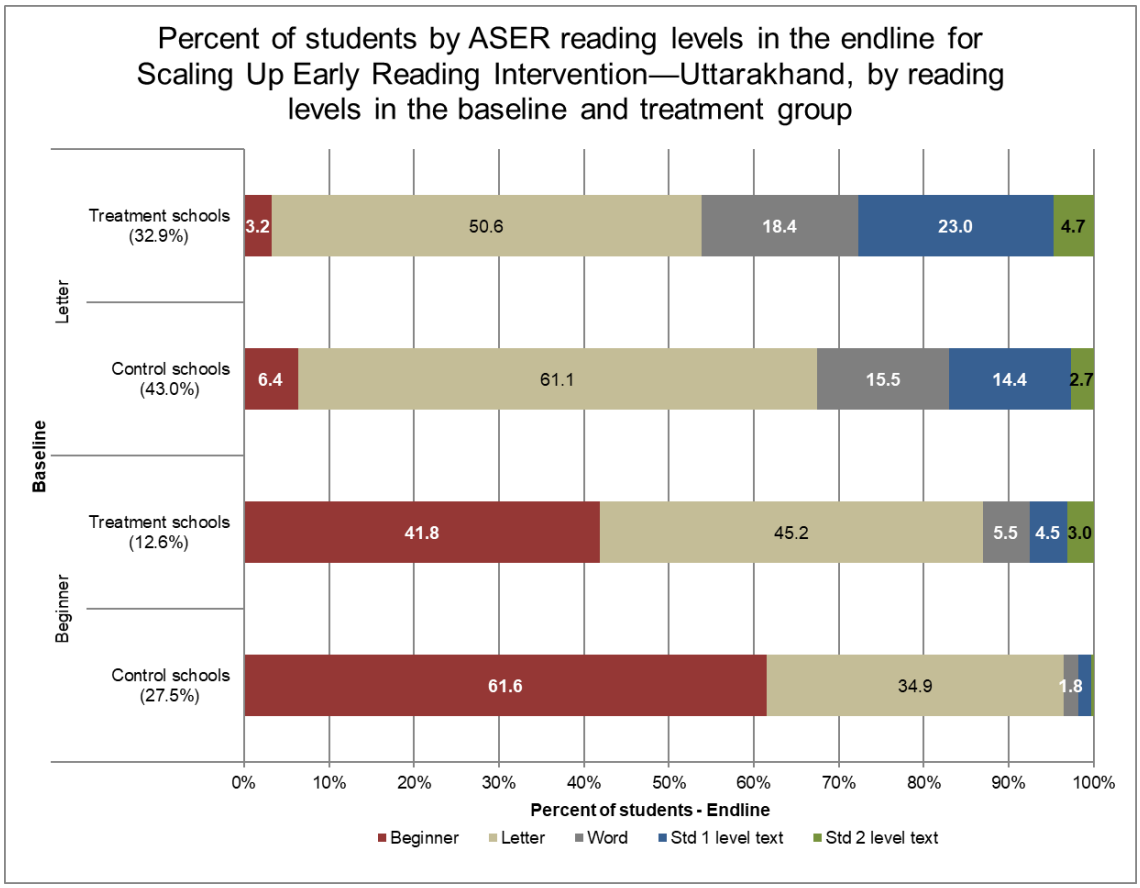
*SEs in parentheses

Another way to understand the effect of the Scaling Up Early Reading Intervention project in Uttarakhand on reading performance is to see how well the program assisted students at lower levels of reading to progress to higher levels.

At the initial assessment, 12.6% of students in treatment schools and 27.5% of students in control schools were marked at the beginner level. From ***Figure 3***, we can see that of these students, in control schools, 61.6% were still at the beginner level, compared to 41.8% of those in treatment schools. In treatment schools, 45.2% of the students who were marked at the beginner level at the initial assessment had progressed to the letter level at the endline assessment, compared to 34.9% of students in control schools.

At the initial assessment, 32.9% of students in treatment schools were marked at the letter level, compared to 43% of students in control schools. Of these students, 61.1% in control schools were still at the letter level at the endline assessment, compared to 50.6% of students in treatment schools. At the endline assessment, 18.4% and 23.0% of students in treatment schools who were categorized at the letter level at the initial assessment had progressed to the word and Standard 1 text levels, respectively. These proportions were 15.5% and 14.4%, respectively, for students in control schools.

Figure 3. Changes in the percentage of students at different ASER reading levels at endline (Scaling Up Early Reading Intervention—Uttarakhand)



(2) EGRA Results

As described in the Executive Summary, the intervention impact was calculated for each project location by testing an individual student at initial assessment and retesting that same student at endline. The difference, typically a gain, in student scores was calculated for each individual student. The individual student gains were then averaged for the treatment group to calculate the mean gain in student scores for that treatment group. This same procedure was repeated for students in the control group. The difference in mean gains between treatment and control is the intervention impact.

Table 9 displays the gain in mean scores and effect size for ORF and reading comprehension for Standard 2 students by treatment group. The difference (T-C) measures the gains achieved by students in the treatment schools over the control schools. A positive number indicates a gain in the treatment group, and a negative number indicates a gain for the control group. The effect size is an indication of intervention effect. The higher the number, the less we can attribute any gains solely to the intervention in question.

For the Scaling Up Early Reading Intervention—Uttarakhand, a positive impact was found that may be partly attributable to the intervention. At endline, the treatment group saw greater gains in ORF means between the initial assessment and endline compared to the control. Treatment schools had a mean gain of 11.5 cwpm, while the control group had a mean gain of 6.1 cwpm; therefore, the treatment group's gain was, on average, 5.4 cwpm higher than that of the control group. This difference is significant at the 0.01 level, as indicated by two asterisks.

Table 9. Average student gains in ORF and reading comprehension scores (Scaling Up Early Reading Intervention—Uttarakhand)

Subtasks	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	6.1 (0.7)	5.4**	0.24 Small
	Treatment	11.5 (0.7)		
Mean gain in reading comprehension (% correct)	Control	5.8% (1%)	7.7%**	0.20 Small
	Treatment	13.5% (1.2%)		

**Significant at 0.01 level

At endline, students in the treatment group had a higher reading comprehension gain than those in the control group. Treatment schools had a mean gain of 13.5%, while the control group had a mean gain of 5.8%; therefore, the treatment group's gain was, on average, 7.7% higher than that of the control group.

Although the treatment group had greater impact on their reading outcomes for both ORF and reading comprehension, how much of that gain is attributable to inherent differences between the treatment and control groups or to any early intervention effect cannot be determined. For example, as documented in the Initial Assessment Report, students in the treatment group were reading better on average than students in the control group.

Figure 4 is a scatterplot of initial student scores (x-axis) and endline scores (y-axis). The size of the circles indicates the frequency of students in the sample with that result. The larger the dot, or bubble, the more students it represents. For example, the largest blue bubble at Point A represents 227 students in the control group who scored 0 at the initial assessment and also scored 0 at endline, whereas the red bubble at Point B represents one student who scored 80 cwpm at the initial assessment and 94 cwpm at endline.

The scatterplot is a helpful visualization of the differences in performance between control and treatment groups. Regression lines were fitted for the treatment and control scatters separately. The treatment line (red) is above the control line (blue) and has a greater slope, confirming that, on average, students in treatment schools experienced greater individual gains.

Students in the treatment group had a higher initial reading fluency and performed better than control students at endline, as indicated by the clustering of red circles. The size of the blue circle at (0,0) indicates that a disproportionate number of students scored 0 cwpm at both the initial assessment and the endline assessment (control). As a point of reference, at the initial assessment, 31% of treatment group students scored zero, compared to 61% of control group students.

Figure 4. Scatterplot of initial and endline assessment scores (Scaling Up Early Reading Intervention—Uttarakhand)

Table 10 displays the mean initial and endline scores by treatment group for ORF and reading comprehension. It is important to note that the scores presented were not used to calculate intervention impact. Instead, the values in ***Table 10*** represent the means of the treatment and control groups at initial assessment and endline without matched individual student scores.

For both the treatment and control groups, the average ORF and reading comprehension scores improved from the initial to endline assessments. However, the improvement in average scores for the treatment group was greater on both measures. Whether the amount of improvement for either treatment or control is adequate for the amount of elapsed instruction time between the initial and endline assessments cannot be determined.

Table 10. Mean ORF and reading comprehension scores at the initial assessment and endline across treatment groups (Scaling Up Early Reading Intervention—Uttarakhand)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	7.3 (1.0)	13.5 (1.4)
	Treatment	15.7 (1.2)	27.2 (1.6)
Reading comprehension (% correct)	Control	7.0% (1.2%)	12.8% (1.7%)
	Treatment	13.7% (1.4%)	27.1% (2.0%)

Table 11 describes the percentage of students scoring zero at initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. For both the ORF and reading comprehension subtasks, significantly fewer students in the treatment group scored zero compared to the control group, which confounds the attribution of greater gains (or reductions in zeros) to the treatment alone. Notably, in both the control and treatment groups, the percentage of students scoring zero decreased between the initial assessment and endline.

Table 11. Percentage of students scoring zero at the initial assessment and endline across treatment groups (Scaling Up Early Reading Intervention—Uttarakhand)

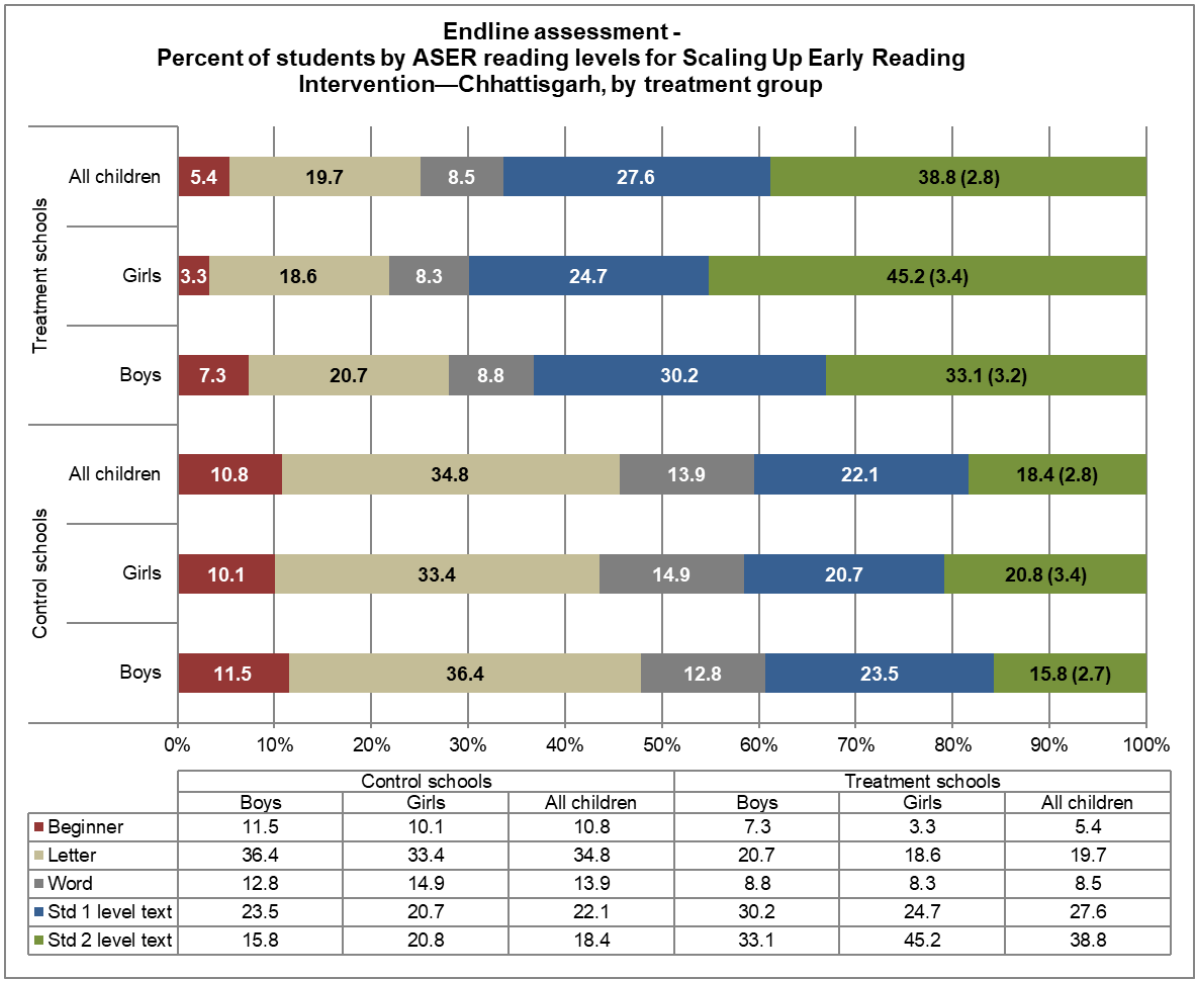
Percentage of Students Scoring Zero	Treatment Group	Initial Assessment (SE)	Endline (SE)
ORF	Control	61.2% (3.4)	39.5% (3.6)
	Treatment	31.1% (2.8)	14% (1.4)
Reading comprehension	Control	81.9% (2.7%)	72.4% (2.9%)
	Treatment	66.6% (3.1%)	46.3% (3%)

4.1.2 Scaling Up Early Reading Intervention Project—Chhattisgarh

(1) ASER Results

Figure 5 shows the performance of Standard 2 students in treatment and control schools at the endline assessment for the Scaling Up Early Reading Intervention project in Chhattisgarh. Of the students in treatment schools, 38.8% could read the Standard 2-level text, and another 27.6% could read the Standard 1-level text but not the Standard 2-level text. In control schools, these proportions were 18.4% and 22.1%, respectively. At the endline assessment, 10.8% of Standard 2 students in control schools could not even read letters, compared to 5.4% of students in treatment schools.

Figure 5. Percentage of students by ASER reading level at endline (Scaling Up Early Reading Intervention Project—Chhattisgarh)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	11.5	10.1	10.8	7.3	3.3	5.4
Letter	36.4	33.4	34.8	20.7	18.6	19.7
Word	12.8	14.9	13.9	8.8	8.3	8.5
Std 1 level text	23.5	20.7	22.1	30.2	24.7	27.6
Std 2 level text	15.8	20.8	18.4	33.1	45.2	38.8

*SEs in parentheses

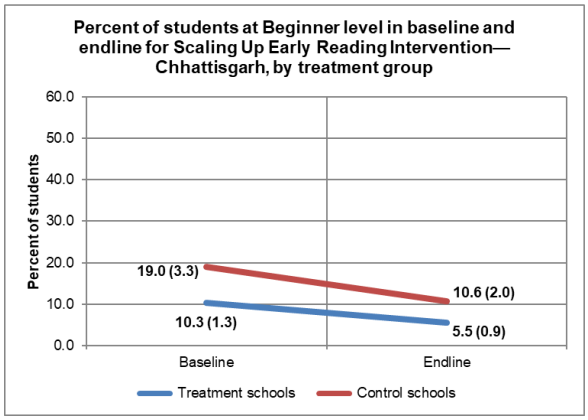
To understand the effect of the Scaling Up Early Reading Intervention project in Chhattisgarh on students' reading performance, we looked at the change in the proportions of students at the beginner and Standard 2 text levels from the initial assessment to the endline assessment for treatment and control schools.

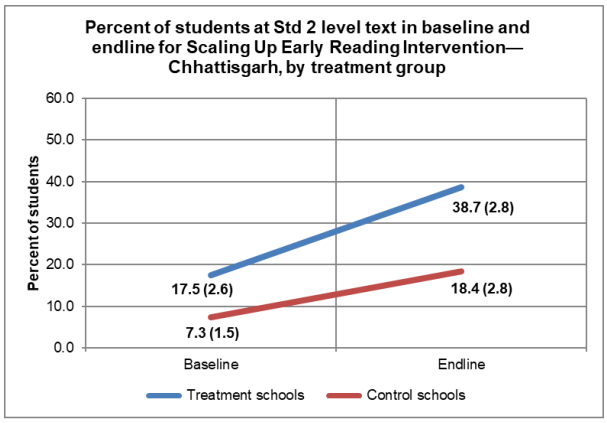
From ***Figure 6***, we can see that the proportion of students at the beginner level decreased from 10.3% at the initial assessment to 5.5% at the endline assessment for treatment schools. For control schools, this proportion reduced from 19% to 10.6%. Regression analysis (see ***Annex H***) confirmed that the change in the proportion of students at the

beginner level from the initial to endline assessments did not differ significantly between treatment (4.8% points) and control schools (8.4% points).

Figure 6 shows that the proportion of students in treatment schools who could read the Standard 2-level text increased from 17.5% at the initial assessment to 38.7% at the endline assessment. The corresponding increase in control schools was 7.3% to 18.4%. Regression analysis (see ***Annex G***) showed that the change in the proportion of students who could read the Standard 2-level text from the initial to endline assessments was significantly larger in treatment schools (21.2% points) than in control schools (11.1% points), confirming the impact of the Scaling Up Early Reading Intervention project in Chhattisgarh.

Figure 6. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (Scaling Up Early Reading Intervention Project—Chhattisgarh)





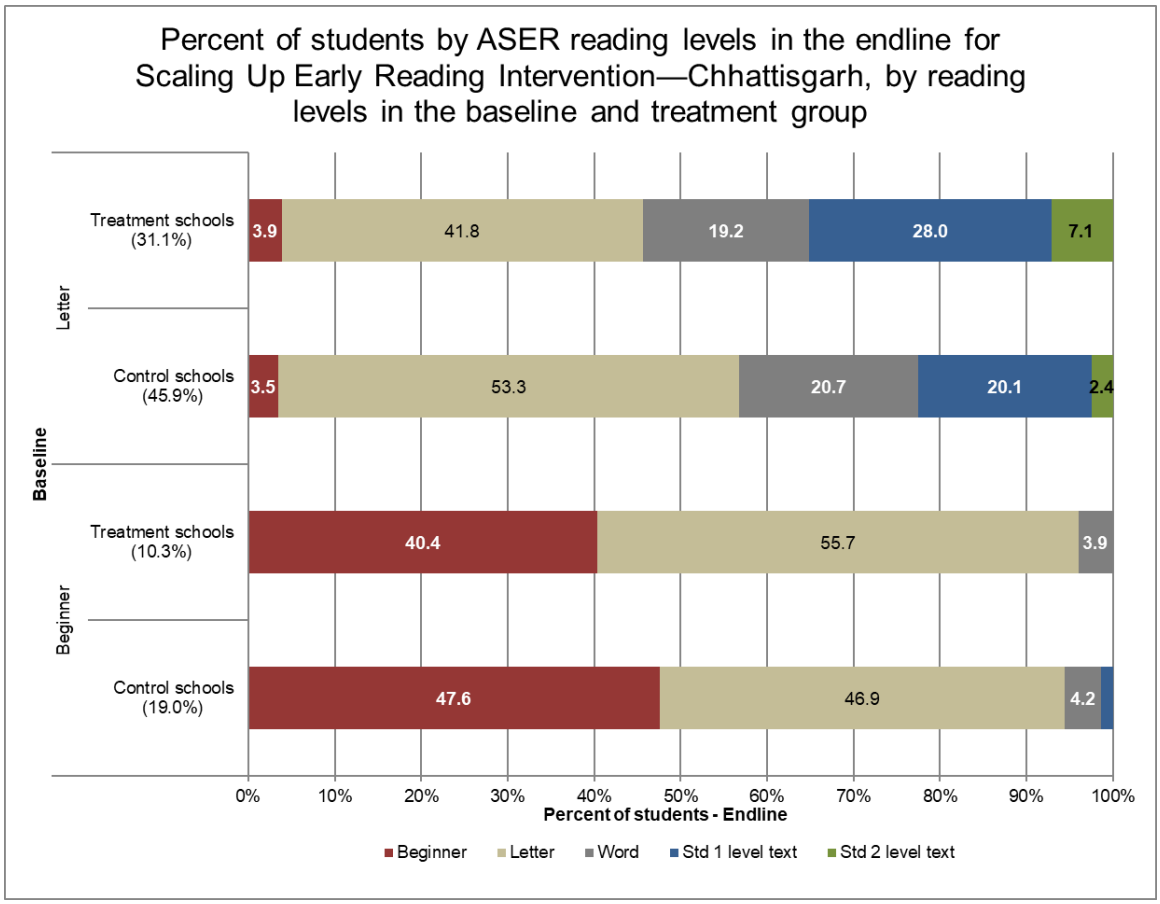
*SEs in parentheses

Another way to understand the effect of the Scaling Up Early Reading Intervention project in Chhattisgarh on reading performance is to see how well the program assisted students at lower levels of reading to progress to higher levels.

At the initial assessment, 10.3% of Standard 2 students in treatment schools and 19% of students in control schools were marked at the beginner level. ***Figure 7*** shows that of these students, 47.6% of those in control schools were still at the beginner level compared to 40.4% of those in treatment schools. In treatment schools, 55.7% students who were marked at the beginner level at the initial assessment had progressed to the letter level at the endline assessment, compared to 46.9% students in control schools.

At the initial assessment, 31.1% of the students in treatment schools were marked at the letter level, compared to 45.9% of students in control schools. Of these students, 53.3% of those in control schools were still at the letter level at the endline assessment, compared to 41.8% of those in treatment schools. More than 35% of the students in treatment schools who were categorized at the letter level at the initial assessment had progressed to Standard 1-level text or higher, compared to approximately 23% of students in the control schools.

Figure 7. Changes in the percentage of students at different ASER reading levels at endline (Scaling Up Early Reading Intervention Project—Chhattisgarh)



(2) EGRA Results

The mean gains in ORF and reading comprehension scores for Scaling Up Early Reading Intervention—Chhattisgarh are shown in ***Table 12***.

At endline, the treatment group demonstrated gains over the control group in both ORF and reading comprehension that are significant at the 0.01 level. The treatment group's mean scores increased on average by 13.6 cwpm between the initial and ednline assessment, as compared to the mean gain of 6.1 cwpm in the control group. The impact of the intervention on the treatment group is that, on average, treatment group students were able to read 7.6 cwpm more compared to the control group.

It is important to take effect size into consideration when interpreting these results. An effect size over 0.2 indicates a small intervention effect, and an effect size approaching 0.5 is considered to reflect a medium intervention effect. Thus, not all gains can be attributed solely to the reading intervention. This finding was also noted the initial assessment, where the treatment group's average ORF was 14.6 cwpm, that of the control group was 6.6 cwpm, and the effect size was 0.57 (***Table 12***). Similar to the project in Uttarakhand, students in the treatment group performed better on average than those in the control group on the initial assessment: over half of the students in the control group scored zero, compared to roughly a quarter of the students in the treatment group.

Table 12. Average student gains in ORF and reading comprehension scores (Scaling Up Early Reading Intervention—Chhattisgarh)

Subtask	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	6.1 (0.6)	7.6**	0.43 Small/medium
	Treatment	13.6 (0.7)		
Mean gain in reading comprehension (% correct)	Control	4.9% (0.8%)	10.2%**	0.33 Small
	Treatment	15.1% (1.1%)		

**Significant at the 0.01 level

Treatment also outperformed control when looking at the mean gains in reading comprehension scores. Reading comprehension was measured by the percentage of questions students answered correctly. Treatment schools had a mean gain of 15.1%, while the control group had a mean gain of 4.9%; therefore, the treatment group's gain was, on average, 10.2% higher than that of the control group (significant at the 0.01 level).

Figure 8 is a scatterplot of student scores at initial assessment (x-axis) and endline (y-axis). The vast majority of students' ORF scores improved between the initial assessment and endline for both treatment groups. Similarly to Uttarakhand, the control group had a larger proportion of children scoring zero. This graph demonstrates that the treatment group was more successful in reducing zero scores at endline, as evidenced by the concentration of red bubbles along the y-axis, which indicate a shift away from zero scores for the treatment group. ***Figure 8*** also shows regression lines for the treatment and control scatters separately. The treatment line (red) is above and has a greater slope than the control line (blue), confirming that, on average, students in treatment schools experienced greater individual gains.

Figure 8. Scatterplot of initial assessment and endline scores (Scaling Up Early Reading Intervention—Chhattisgarh)

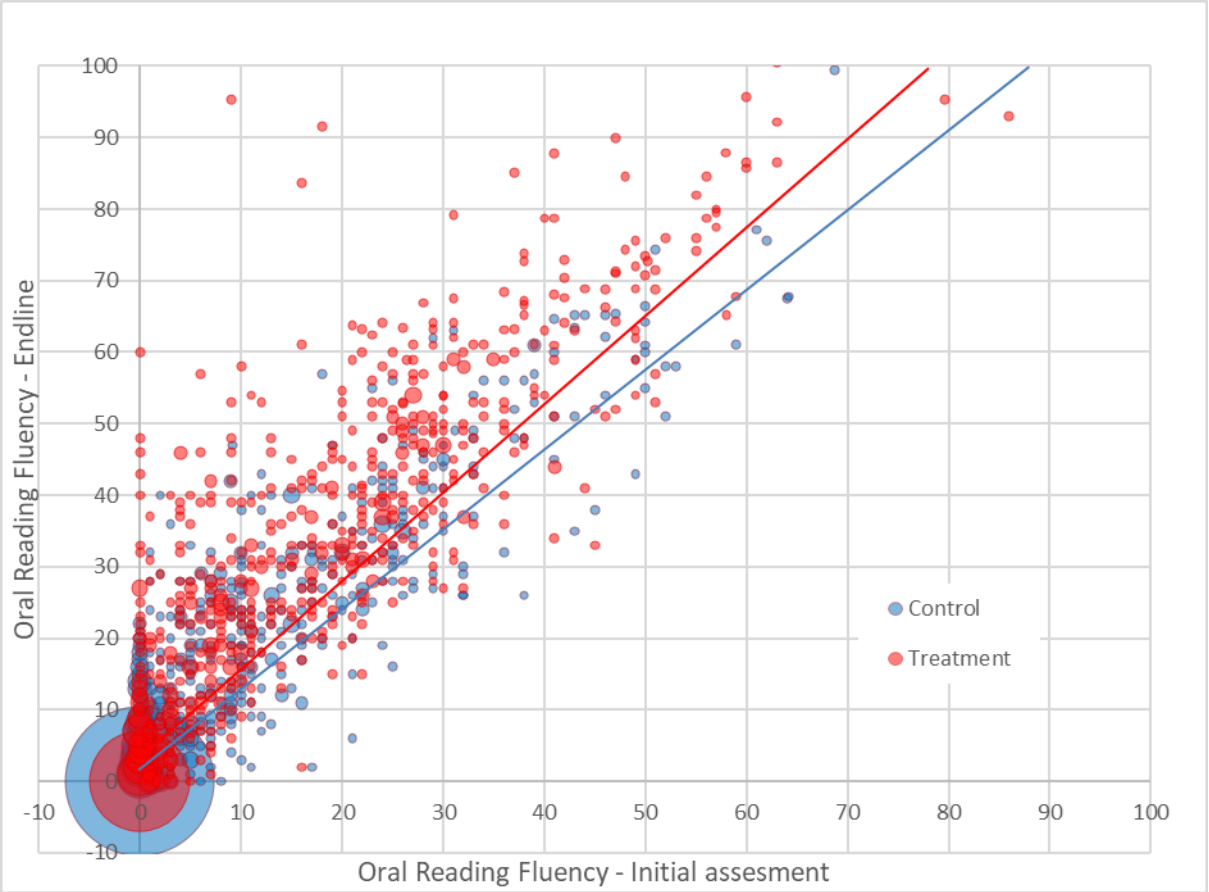


Table 13 displays the initial and endline mean scores ORF and reading comprehension scores by treatment group. It is important to note that these scores were not used to calculate the intervention impact. The figures represent the means of the treatment and control groups at the initial assessment and endline without matched individual student scores.

For both the treatment and control groups, the average ORF and reading comprehension scores improved from the initial to endline assessments. However, the improvement in average scores in the treatment group was greater for both measures. Whether the amount of improvement for either the treatment or control group is adequate for the amount of elapsed instruction time between the initial and endline assessments cannot be determined.

Table 13. Mean ORF and reading comprehension scores at the initial assessment and endline across treatment groups (Scaling Up Early Reading Intervention—Chhattisgarh)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	6.6 (1)	13.0 (1.4)
	Treatment	14.6 (0.9)	28.5 (1.5)
Reading comprehension (% correct)	Control	3.9% (0.8%)	8.9% (1.3%)
	Treatment	9.8% (0.9%)	25.5% (1.7%)

Table 14 describes the percentages of students scoring zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. For both the ORF and reading comprehension subtasks, significantly fewer students in the treatment group scored zero compared to the control group. Notably, in both the control and treatment groups, the percentage of students scoring zero decreased between the initial assessment and endline.

Table 14. Zero scores at the initial assessment and endline across treatment groups (Scaling Up Early Reading Intervention—Chhattisgarh)

Percentage of Students Scoring Zero	Treatment Group	Initial assessment & SE	Endline & SE
ORF	Control	52% (4.5%)	33.3% (4%)
	Treatment	27.1% (2.5%)	13.4% (1.6%)
Reading comprehension	Control	87.8% (2.2)	75.3% (3.1)
	Treatment	69.7% (2.1)	44.4% (2.6)

4.2 Nurturing Early Literacy

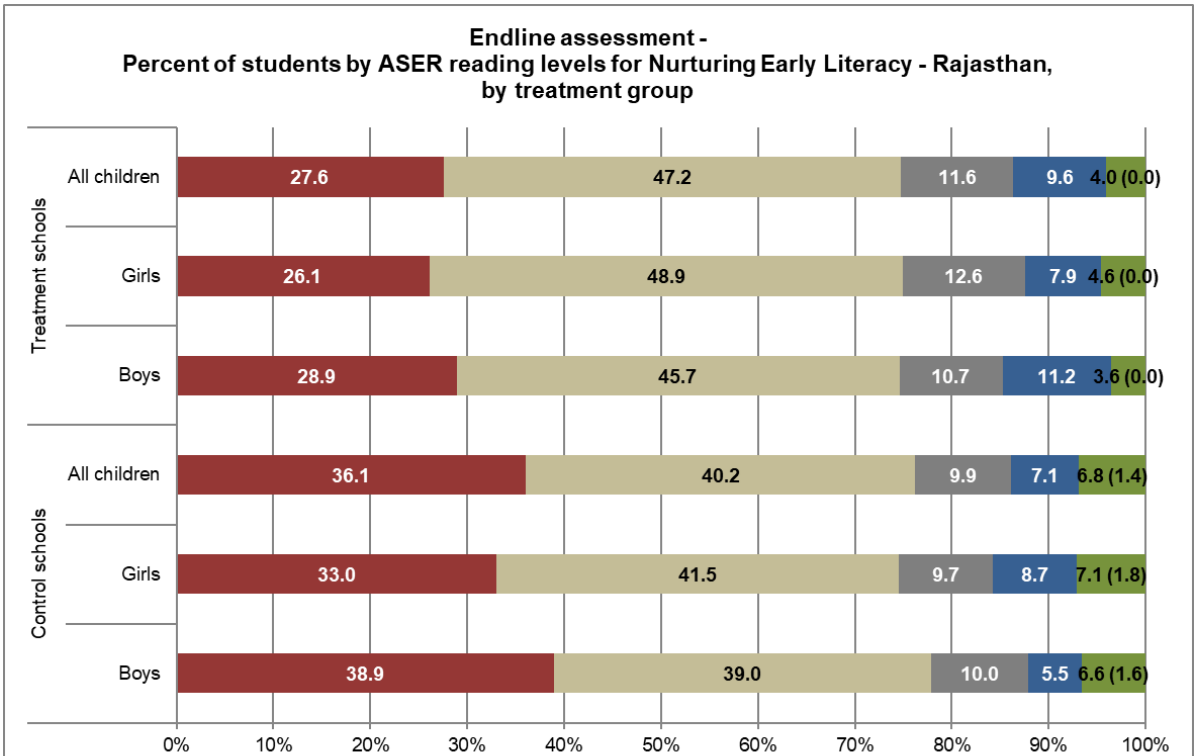
Nurturing Early Literacy is a project co-funded by USAID/India and Tata Trusts, implemented by the CmF, and supported by field-level partners in each state. The overall goal of the project is to build a strong foundation for emergent and early literacy competencies for more than 93,000 students in Standards 1 to 7 in select blocks in three states: Rajasthan, Maharashtra, and Karnataka. The project is currently in the second year of implementation. It began in September 2015 and will continue through September 2019. The areas where the project is being implemented were selected based on an analysis of gaps in education access, delivery, pedagogy, and learning outcomes, along with socioeconomic indicators. The districts selected are Sirohi in Rajasthan, Satara in Maharashtra, and Yadgir in Karnataka. The assessments will be conducted in three languages: Hindi, Kannada, and Marathi. In Rajasthan, only those schools implemented by Bodh Shiksha Samiti were selected for the initial assessment.

4.2.1 Nurturing Early Literacy—Rajasthan

(1) ASER Results

Figure 9 shows the performance of Standard 2 students on the endline assessment in treatment and control schools for the Nurturing Early Literacy project in Rajasthan. Of the students in treatment schools, 27.6% could not even read letters at the endline assessment. In comparison, this proportion was 36.1% in control schools. Approximately 14% of students in both the treatment and control schools could read the Standard 1-level text or higher (13.6% and 13.9%, respectively).

Figure 9. Percentage of students by ASER reading level at endline (Nurturing Early Literacy—Rajasthan)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	38.9	33.0	36.1	28.9	26.1	27.6
Letter	39.0	41.5	40.2	45.7	48.9	47.2
Word	10.0	9.7	9.9	10.7	12.6	11.6
Std 1 level text	5.5	8.7	7.1	11.2	7.9	9.6
Std 2 level text	6.6	7.1	6.8	3.6	4.6	4.0

*SEs in parentheses

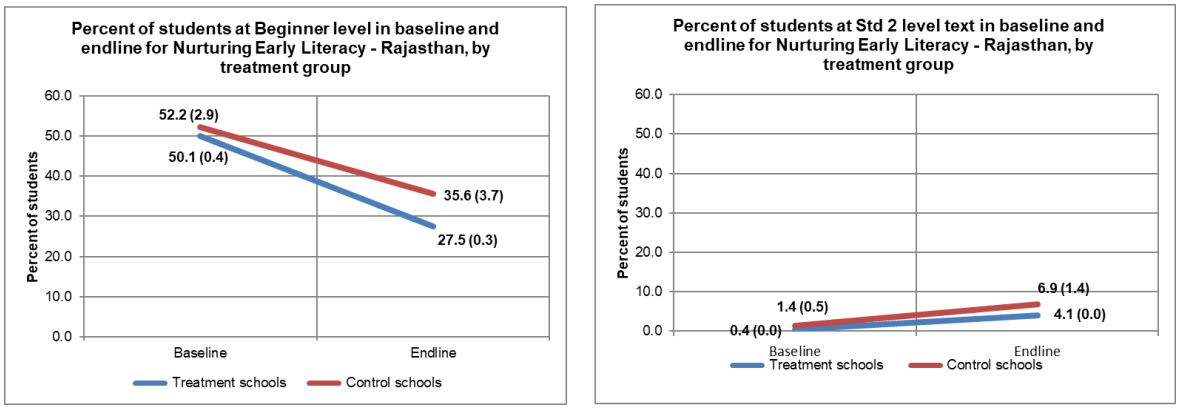
To understand the effect of the Nurturing Early Literacy project in Rajasthan on students' reading performance, we examined the change in the proportion of students in treatment and control schools at the beginner and Standard 2 text levels from the initial assessment to the endline assessment.

From ***Figure 10***, we can see that the proportion of students in treatment schools at the beginner level decreased from 50.1% at the initial assessment to 27.5% at the endline assessment. For control schools, this proportion decreased from 52.2% to 35.6%. Regression analysis (see ***Annex H***) confirmed that the change in the proportion of students

at the beginner level from the initial to endline assessments did not differ significantly between treatment (22.7% points) and control schools (16.6% points).

Figure 10 also shows that the proportion of students who could read the Standard 2-level text increased marginally from 0.4% at the initial assessment to 4.1% at the endline assessment for treatment schools. In control schools, the proportion increased from 1.4% to 6.9%. The change in the proportion of students who could read the Standard 2-level text from the initial to endline assessments did not differ significantly between treatment (3.7% points) and control schools (5.5% points) (see ***Annex G***).

Figure 10. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (Nurturing Early Literacy—Rajasthan)



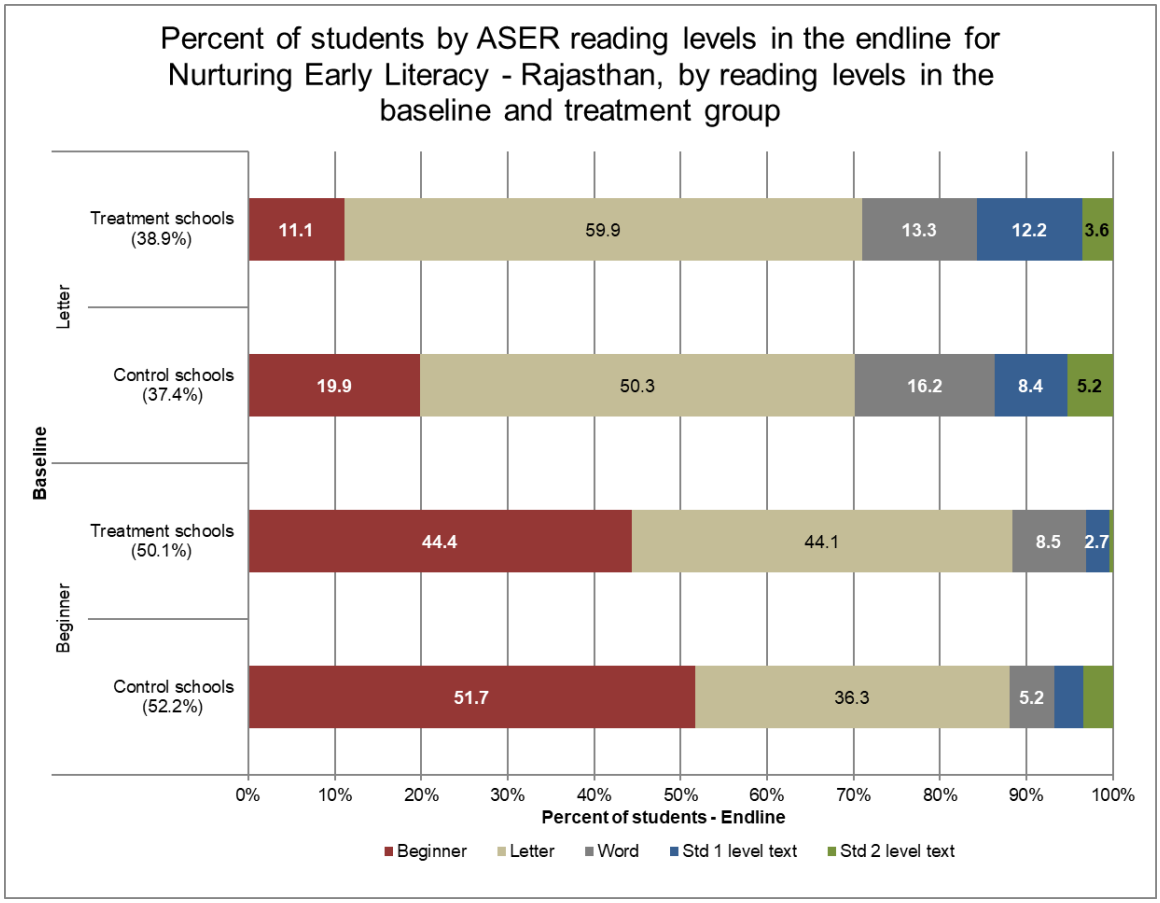
*SEs in parentheses

Another way to understand the effect of the Nurturing Early Literacy project in Rajasthan on reading performance is to see how well the program assisted students at lower levels of reading to progress to higher levels.

At the initial assessment, more than 50% of students (50.1% in treatment schools and 52.2% in control schools) could not even read letters. From ***Figure 11***, we can see that of the students who were marked at the beginner level, 51.7% were still at the beginner level in control schools, compared to 44.4% of students in treatment schools. In treatment schools, 44.1% of students who were marked at the beginner level at the initial assessment had progressed to the letter level at the endline assessment, compared to 36.3% of students in control schools.

At the initial assessment, 38.9% and 37.4% of students in treatment and control schools, respectively, were marked at the letter level. ***Figure 11*** shows that 59.9% of students in treatment schools who were categorized at the letter level at the initial assessment were still at the letter level in the endline assessment, compared to 50.3% of students in control schools. However, in control schools, almost 20% of the letter-level students regressed to the beginner level between the initial and endline assessments. In comparison, this proportion is 11.1% for treatment schools. For both treatment and control schools, approximately 29% of students who were marked at the letter level at the initial assessment had progressed to the word level or higher.

Figure 11. Change in the percentage of students at different ASER reading levels at endline (Nurturing Early Literacy—Rajasthan)



(2) EGRA Results

For Nurturing Early Literacy—Rajasthan, both the treatment and control groups experienced mean gains between the initial assessment and endline (***Table 15***). However, no statistically significant impact on reading outcomes was found as a result of the intervention.

The treatment group's mean scores increased on average by 4.0 cwpm between the assessment and endline, as compared to the mean gain of 5.1 cwpm in the control group. The Initial Assessment Report found that a higher percentage of students were categorized as having low SES in treatment schools (57% in control schools and 69% in treatment schools) and that a higher percentage of students in control schools fell into the high-SES category; that is, 17% of control students were categorized as having high SES, as compared with 5% of students in treatment schools. However, at the initial assessment, performance was nearly identical across the control and treatment groups, with students in control schools reading 1.3 cwpm and those in treatment schools reading 1.7 cwpm. At endline, student scores showed a small improvement.

Table 15. Average student gains in ORF and reading comprehension scores (Nurturing Early Literacy—Rajasthan)

	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	5.1 (0.8)	-1.13	-0.02
	Treatment	4.0 (0)		
Mean gain in reading comprehension (% correct)	Control	3.95% (0%)	-0.84%	-0.01
	Treatment	3.11% (0%)		

Control outperformed treatment in terms of the mean gains in reading comprehension scores. Looking at the impact at endline for both the treatment and control groups, on average, students improved very little, with no statistically discernable difference in the gains achieved in treatment and control schools.

Figure 12 is a scatterplot of student scores at the initial (x-axis) and endline assessments (y-axis). A large proportion of students scoring zero remained scoring zero at endline across both treatment groups. Although small positive gains in the means for both groups were found, the scores are generally clustered at the lower, left-hand quadrant, with few children reading at the benchmark (***Section 5***), and the majority remaining at 0 cwpm. Although the control line (blue) is slightly above the treatment line (red), the slope confirms that there was no statistically discernable difference in the gains achieved in treatment and control schools.

Figure 12. Scatterplot of initial assessment and endline scores (Nurturing Early Literacy—Rajasthan)

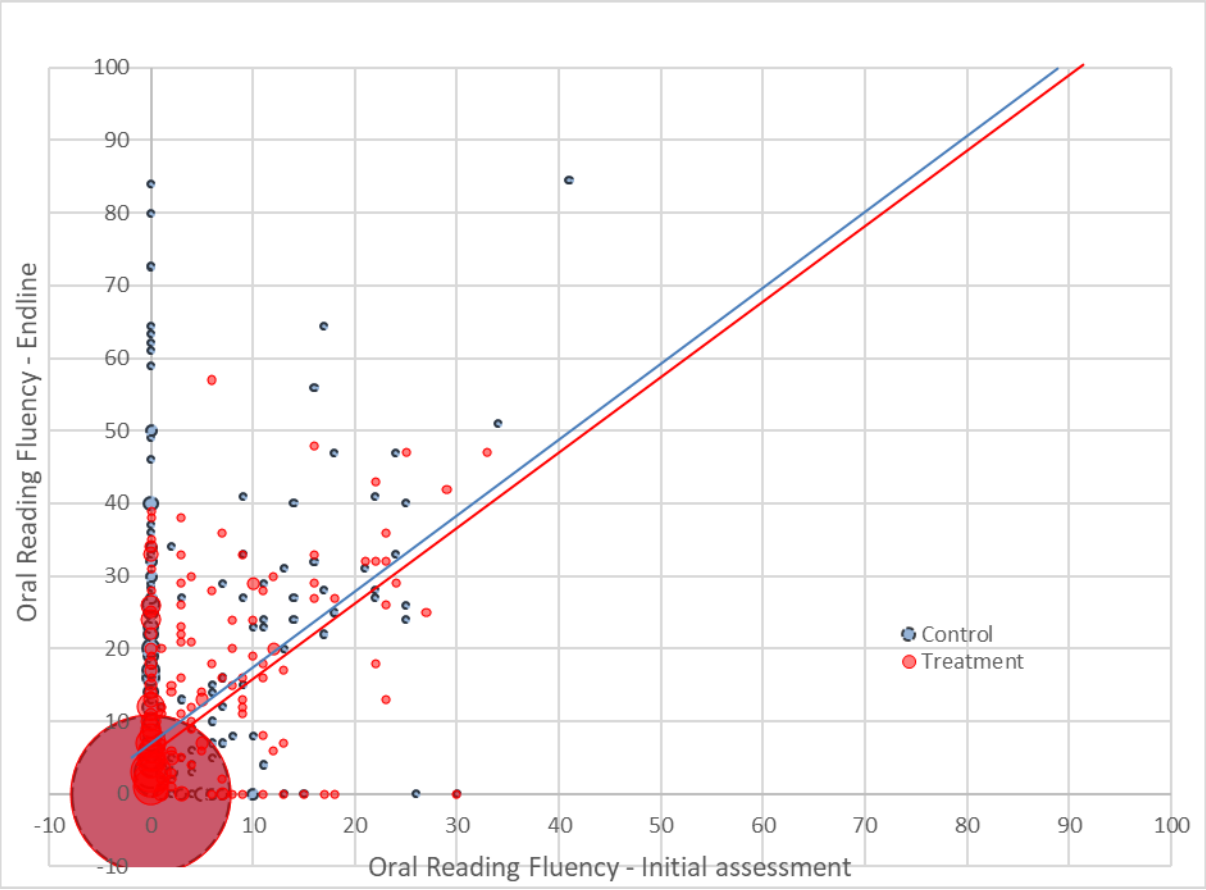


Table 16 displays the mean ORF and reading comprehension scores recorded at the initial assessment and endline for the treatment and control groups. It is important to note that the scores were not used to calculate the intervention impact. These figures represent the mean values calculated for the treatment and control groups at the initial assessment and endline without matched individual student scores.

Table 16 shows how low the scores were for both treatment and control scores at the initial assessment and the small amounts of improvement at endline. The mean ORF scores for both treatment groups were less than 2 cwpm at the initial assessment and improved to approximately 6 cwpm at endline. For reading comprehension, students in both the treatment and control groups achieved less than 1% correct on average at the initial assessment and less than 5% at endline.

Table 16. ORF and reading comprehension means at the initial assessment and endline across treatment groups (Nurturing Early Literacy—Rajasthan)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	1.3 (0.3)	6.4 (0.9)
	Treatment	1.7 (0)	5.7 (0)
Reading comprehension (% correct)	Control	0.7% (0.2%)	4.6% (0.9%)
	Treatment	0.6% (0%)	3.8% (0%)

Table 17 describes the percentage of students scoring zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. In both groups, the percentage of students scoring zero on reading comprehension and ORF decreased. However, the observed reductions did not differ significantly between the treatment and control groups.

Table 17. Zero scores at the initial assessment and endline across treatment groups (Nurturing Early Literacy—Rajasthan)

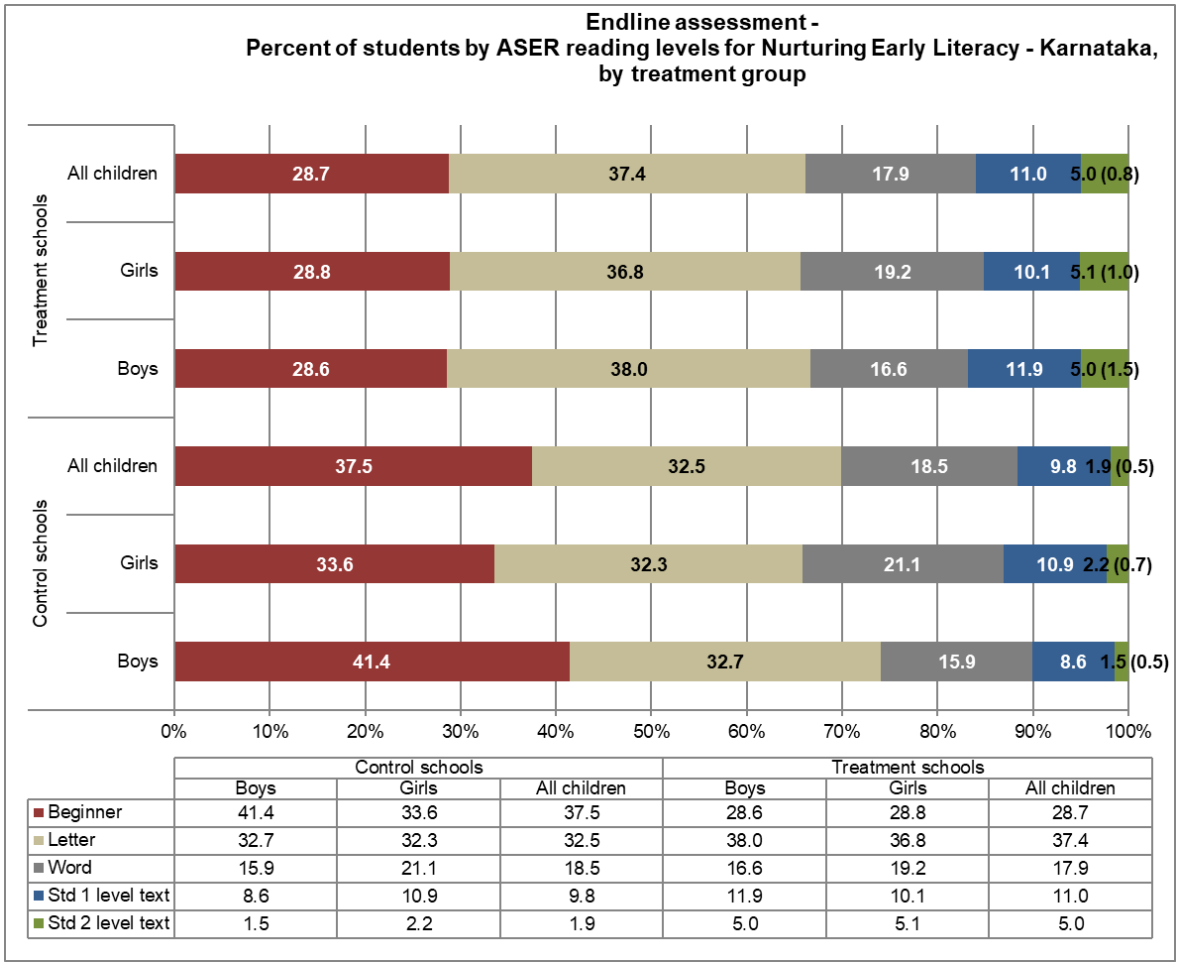
Percentage of Students Scoring Zero	Treatment Group	Initial Assessment (SE)	Endline (SE)
ORF	Control	86.8% (2.3%)	65.2% (3.5%)
	Treatment	80.9% (0.2%)	58.5% (0.4%)
Reading comprehension	Control	97.7% (0.6%)	88.1% (1.9%)
	Treatment	97.2% (0%)	88.1% (0.2%)

4.2.2 Nurturing Early Literacy—Karnataka

(1) ASER Results

Figure 13 shows the performance of Standard 2 students in treatment and control schools at the endline assessment for the Nurturing Early Literacy project in Karnataka. In treatment schools, 28.7% of students could not read even letters at the endline assessment, compared to 37.5% of students in control schools. Additionally, 1.9% of students in control schools could read the Standard 2-level text, compared to 5% of students in treatment schools.

Figure 13. Percentage of students by ASER reading level at endline (Nurturing Early Literacy—Karnataka)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	41.4	33.6	37.5	28.6	28.8	28.7
Letter	32.7	32.3	32.5	38.0	36.8	37.4
Word	15.9	21.1	18.5	16.6	19.2	17.9
Std 1 level text	8.6	10.9	9.8	11.9	10.1	11.0
Std 2 level text	1.5	2.2	1.9	5.0	5.1	5.0

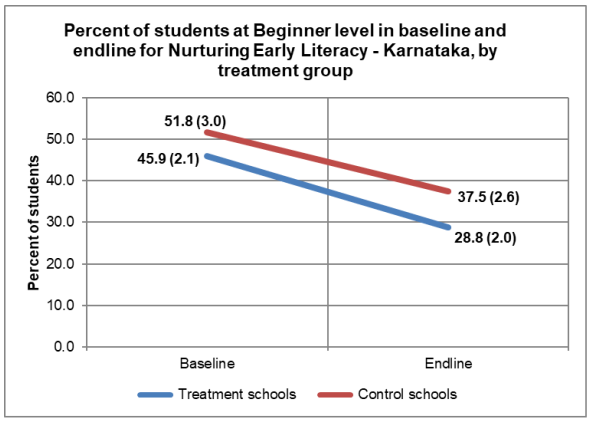
*SEs in parentheses

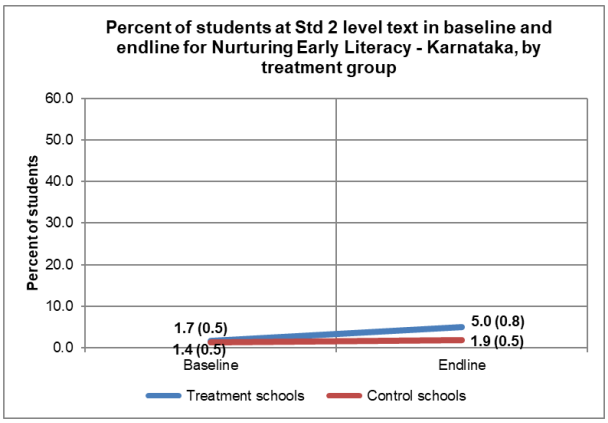
To understand the effect of the Nurturing Early Literacy project in Karnataka on students' reading performance, we looked at the change in the proportion of students at the beginner and Standard 2 text levels from the initial assessment to the endline assessment in treatment and control schools.

From ***Figure 14***, we can see that the proportion of students in treatment schools at the beginner level decreased from 45.9% at the initial assessment to 28.8% at the endline assessment. For control schools, this proportion reduced from 51.8% to 37.5%. Regression analysis (see ***Annex H***) confirmed that the change in the proportion of students at the beginner level from the initial to endline assessments did not differ significantly between treatment (17.1% points) and control schools (14.3% points).

The proportion of students in treatment schools who could read the Standard 2-level text increased marginally from 1.7% at the initial assessment to 5% at the endline assessment. This increase was from 1.4% to 1.9% in control schools. Regression analysis (see ***Annex G***) showed that the in the proportion of students who could read the Standard 2-level text from the initial assessment to endline assessments was significantly greater in treatment schools (3.3% points) than in control schools (0.5% points), confirming the impact of the Nurturing Early Literacy project in Karnataka.

Figure 14. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (Nurturing Early Literacy—Karnataka)





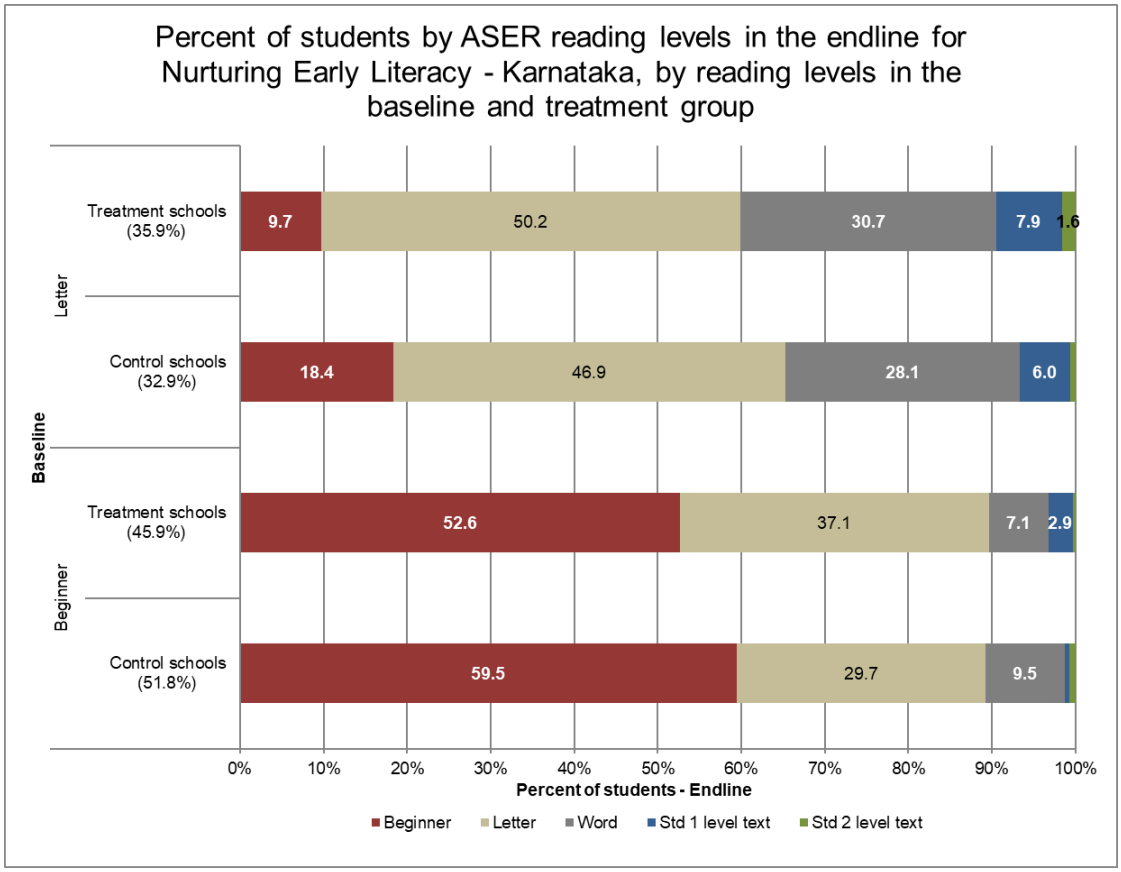
*SEs in parentheses

Another way to understand the effectiveness of the Nurturing Early Literacy project in Karnataka on reading performance is to see how well the program assisted students at lower levels of reading to progress to higher levels.

At the initial assessment, 45.9% and 51.8% of students in treatment and control schools, respectively, could not even read letters. From ***Figure 15***, we can see that of those students who were marked at the beginner level, 59.5% of students in control schools are still at the beginner level, compared to 52.6% of students in treatment schools. In treatment schools, 37.1% students who were marked at the beginner level at the initial assessment had progressed to the letter level at the endline assessment, compared to 29.7% of students in control schools.

At the initial assessment, 35.9% of students in treatment schools and 32.9% of students in control schools were marked at the letter level. ***Figure 15*** shows that 50.2% of the students in treatment schools who were marked at the letter level at the initial assessment were still at the letter level at the endline assessment, compared to 46.9% of students in control schools. However, in control schools, more than 18% of the letter-level students regressed to the beginner level between the initial and endline assessments. In comparison, this proportion is less than 10% for treatment schools. In treatment schools, approximately 40% of students who were marked at the letter level at the initial assessment had progressed to the word level or higher; this proportion is roughly 35% for control schools.

Figure 15. Changes in the percentage of students at different ASER reading levels at endline (Nurturing Early Literacy—Karnataka)



(2) EGRA Results

For Nurturing Early Literacy—Karnataka, both the treatment and control groups experienced small mean gains between initial assessment and endline (***Table 18***). However, no statistically significant impact on reading outcomes was found as a result of the intervention.

Treatment schools had a mean gain of 3.1 cwpm, while the control group had a mean gain of 2.1 cwpm; therefore, the treatment group's gain was, on average, 1.0 cwpm higher than that of the control group. Students in the treatment group had a higher reading comprehension gain compared to the control group. Control schools had a mean gain of 2.3%, while the treatment group had a mean gain of 2.8%; therefore, the treatment group's gain was, on average, 0.5% higher than that of the control group.

Table 18. Average student gains in ORF and reading comprehension scores (Nurturing Early Literacy—Karnataka)

Mean Gain in Score	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	2.1 (0.4)	1.0	0.04
	Treatment	3.1 (0.3)		
Mean gain in reading comprehension (% correct)	Control	2.3% (0.01%)	0.5%	0.01
	Treatment	2.8% (0%)		

As shown in ***Figure 16***, a large proportion of students scoring zero remained scoring zero at endline across both treatment groups. Although some positive gains in means were found for both groups, the scores are generally clustered at the lower, left-hand quadrant, with few children reading at benchmark (***Section 5***) and the majority reading at less than 6 cwpm.

Figure 16. Scatterplot of initial assessment and endline scores (Nurturing Early Literacy—Karnataka)

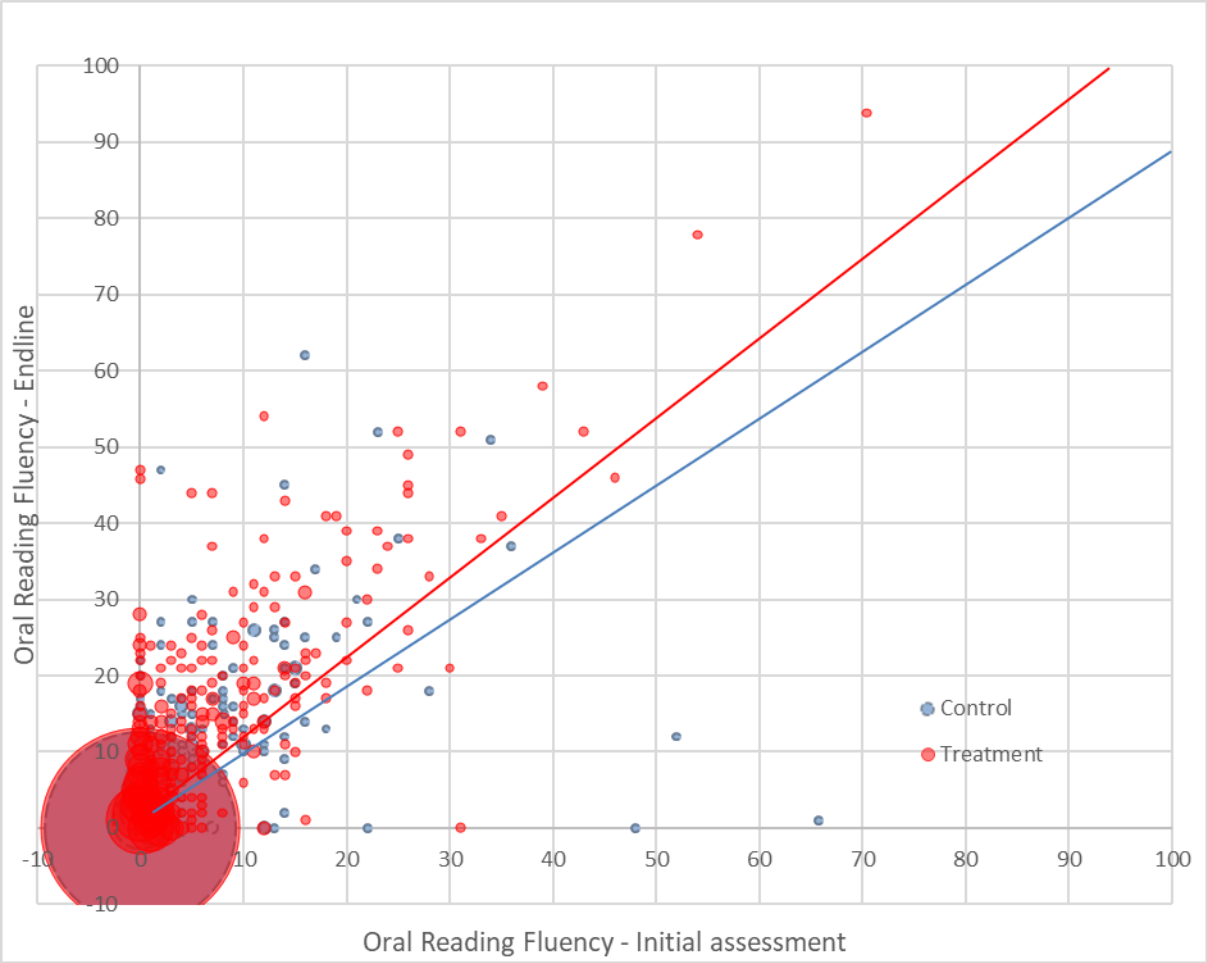


Table 19 displays the mean ORF and reading comprehension scores by treatment group at the initial and endline assessments. It is important to note that the scores were not used to calculate the intervention impact. They represent the means of the treatment and control groups at the initial assessment and endline without matched individual student scores.

For both the treatment and control groups, reading performance remained low at endline, with limited improvement, on average, in both ORF and reading comprehension.

Table 19. Mean ORF and reading comprehension scores at the initial assessment and endline across treatment groups (Nurturing Early Literacy—Karnataka)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	2.1 (0.4)	4 (0.7)
	Treatment	2.6 (0.3)	5.8 (0.5)
Reading comprehension (% correct)	Control	1.0% (0.3%)	3.3% (0.8%)
	Treatment	2.1% (0.4%)	4.9% (0.7%)

Table 20 describes the percentage of students scoring zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. In both groups, the percentage of students scoring zero on reading comprehension and ORF decreased. However, the reductions observed did not differ significantly between the treatment and control groups.

Table 20. Percentage of children scoring zero at the initial assessment and endline across treatment groups (Nurturing Early Literacy—Karnataka)

Percentage of Students Scoring Zero	Treatment Group	Initial Assessment (SE)	Endline (SE)
ORF	Control	71.9% (3.1)	58.1% (3.2)
	Treatment	65.9% (2.3)	50.7% (2.5)
Reading comprehension	Control	96.6% (0.8%)	89.7% (2.2%)
	Treatment	93.3% (1.2%)	84.3% (1.6%)

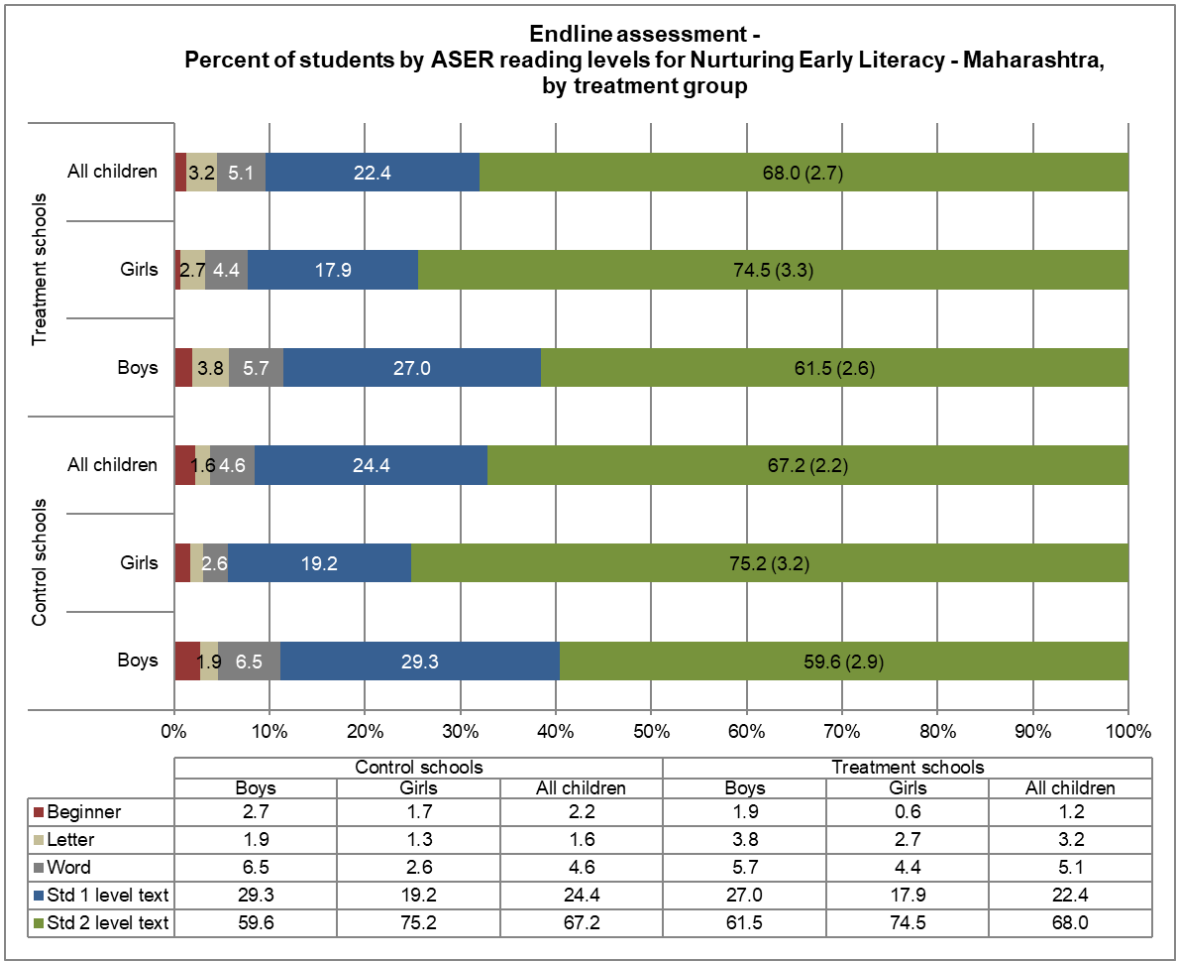
4.2.3 Nurturing Early Literacy—Maharashtra

(1) ASER Results

Figure 17 shows the performance of Standard 2 students in treatment and control schools at the endline assessment for the Nurturing Early Literacy project in Maharashtra. This project had the highest proportion of students reading Standard 2 level text in the initial assessment. This trend is visible even in the endline assessment.

Of the students in treatment schools, 68% could read the Standard 2-level text, whereas in control schools, this proportion was 67.2%. Additionally, 22.4% of students in treatment schools and 24.4% of students in control schools could read the Standard 1-level text but not the Standard 2-level text.

Figure 17. Percentage of students by ASER reading level at endline (Nurturing Early Literacy—Maharashtra)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	2.7	1.7	2.2	1.9	0.6	1.2
Letter	1.9	1.3	1.6	3.8	2.7	3.2
Word	6.5	2.6	4.6	5.7	4.4	5.1
Std 1 level text	29.3	19.2	24.4	27.0	17.9	22.4
Std 2 level text	59.6	75.2	67.2	61.5	74.5	68.0

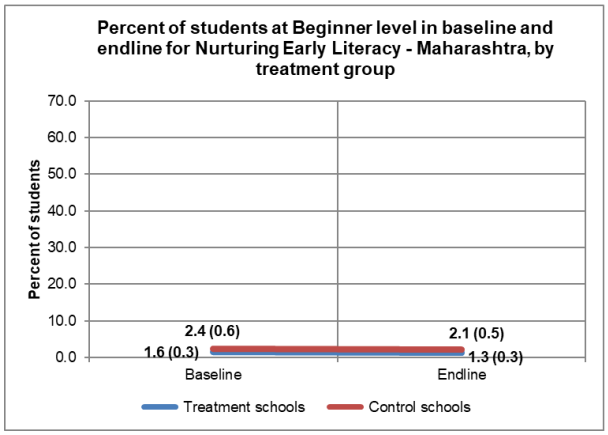
*SEs in parentheses

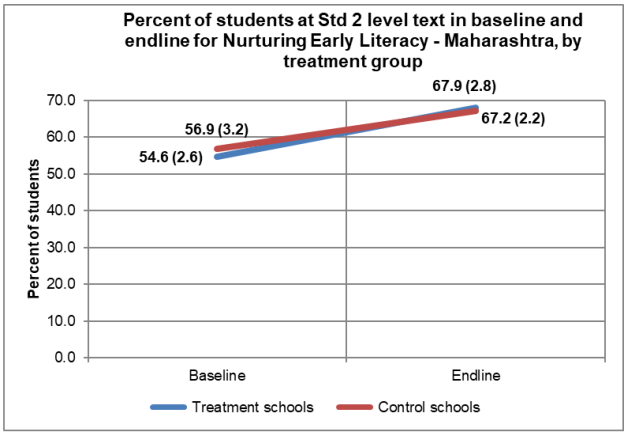
To understand the effect of the Nurturing Early Literacy project in Maharashtra on students' reading performance, we examined the change in the proportion of students at the beginner and Standard 2 text levels in treatment and control schools between the initial assessment and the endline assessment.

From ***Figure 18***, we can see that the proportion of students at the beginner level was low in both treatment and control schools, even at the initial assessment. This proportion remained low and unchanged for both treatment and control schools at the endline assessment without any significant difference.

For treatment schools, the proportion of students who could read the Standard 2-level text increased from 54.6% at the initial assessment to 67.9% at the endline assessment. In control schools, the increase was from 56.9% to 67.2%. The change in the proportion of students who could read the Standard 2-level text from the initial to endline assessments did not differ significantly in treatment (13.3% points) and control schools (10.3% points) (see ***Annex G***).

Figure 18. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (Nurturing Early Literacy—Maharashtra)





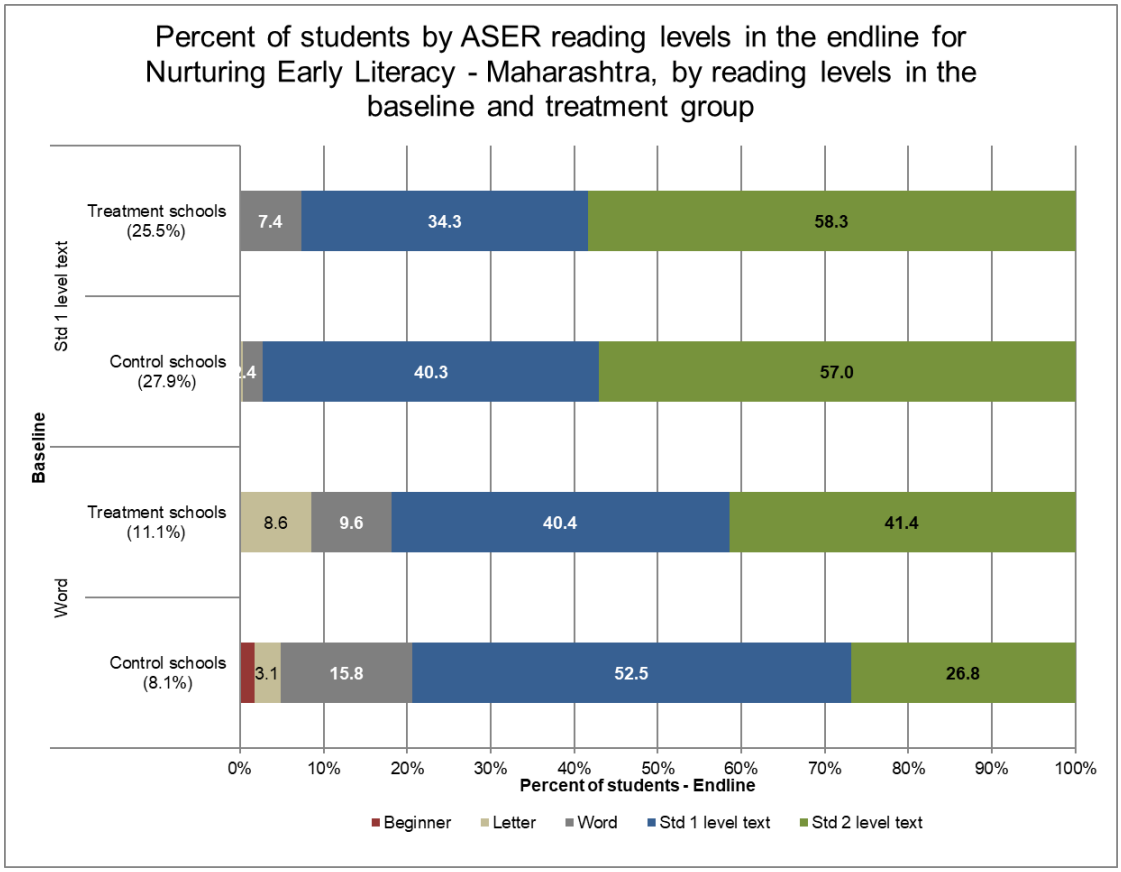
*SEs in parentheses

In the Nurturing Early Literacy project in Maharashtra, a large proportion of students were already reading at the word level or higher at the initial assessment. Therefore, to understand the effect of the Nurturing Early Literacy project in Maharashtra on reading performance, we looked at how well the program assisted students at these levels to progress to the highest level: the Standard 2 text level.

At the initial assessment, 11.1% and 8.1% of students in treatment and control schools, respectively, were marked at the word level. ***Figure 19*** shows that of these students, 41.4% of those in treatment schools had progressed to the Standard 2-level text at the endline assessment. In control schools, this proportion was 26.8%.

At the initial assessment, 25.5% of students in treatment schools and 27.9% of students in control schools were marked at the Standard 1 text level. From ***Figure 19***, we can see that 58.3% of the students in treatment schools who were marked at the Standard 1 text level at the initial assessment had progressed to the Standard 2 text level at the endline assessment. For control schools, this proportion was 57%.

Figure 19. Changes in the percentage of students at different ASER reading levels at endline (Nurturing Early Literacy—Maharashtra)



(2) EGRA Results

For Nurturing Early Literacy—Maharashtra, the control group saw greater gains in mean ORF scores between the initial assessment and endline (***Table 21***) than the treatment group. Control schools had a mean gain of 13.9 cwpm, while the treatment group had a mean gain of 11.9 cwpm; therefore, the control group's gain was, on average, 2 cwpm higher than that of the treatment group. This difference is significant at the 0.5 level.

For reading comprehension, both the treatment and control groups exhibited gains in reading outcomes, but the difference in gains between the groups was not significant. Control schools had a mean gain of 21.1%, while the treatment group had a mean gain of 19.1%; therefore, the control group's gain was, on average, 2% higher than that of the treatment group.

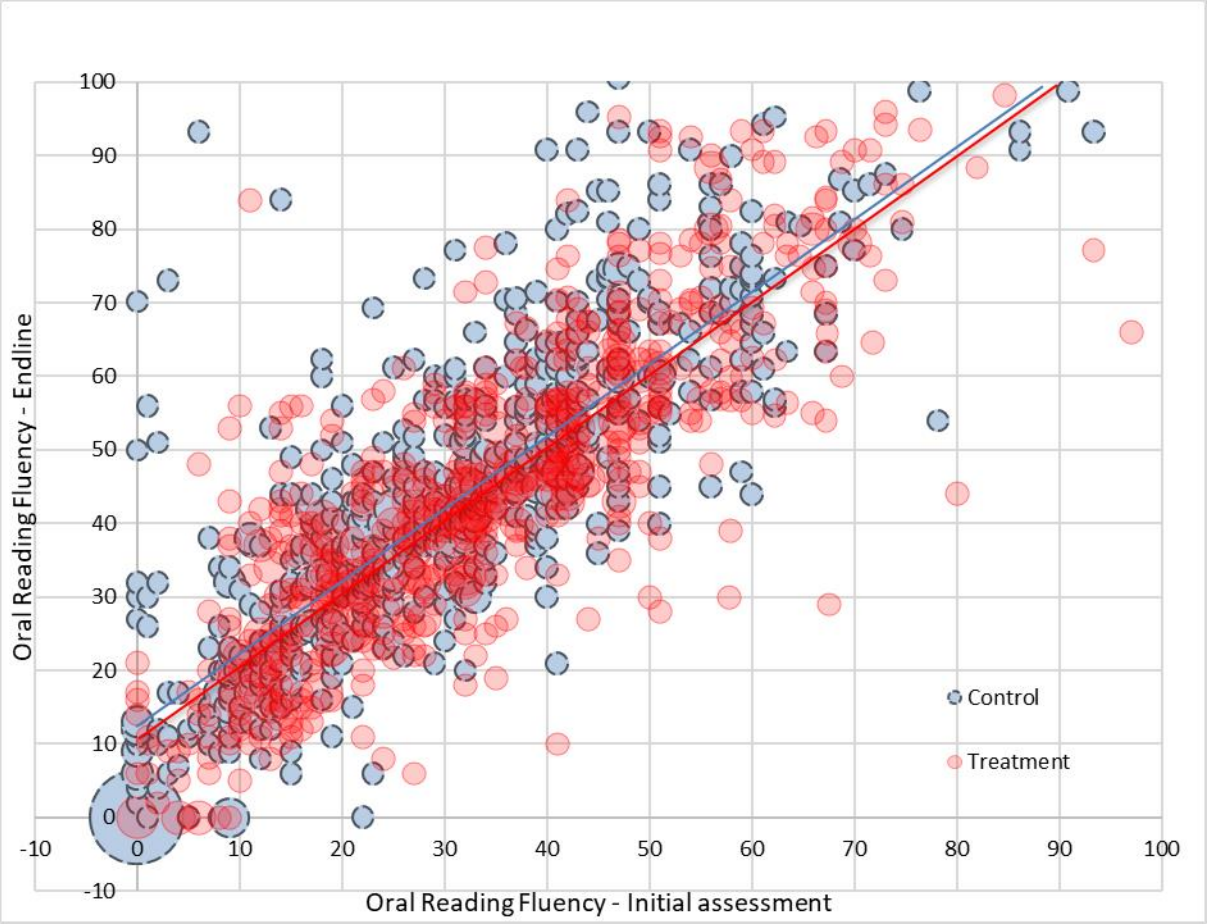
Table 21. Average student gains in ORF and reading comprehension scores (Nurturing Early Literacy—Maharashtra)

Mean Gain in Score	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	13.9 (0.8)	-2.0*	-0.04
	Treatment	11.9 (0.5)		
Mean gain in reading comprehension (% correct)	Control	21.1% (0.0%)	-2.0%	-0.02
	Treatment	19.1% (0.0%)		

*Significant at the 0.5 level

Figure 20 is a scatterplot of student scores at initial assessment (x-axis) and endline (y-axis). Compared to the other CmF project locations, a smaller proportion of students in Maharashtra scored zero. The scatterplot shows that students demonstrated a wide range of reading levels in both treatment and control schools. The scores are not clustered near zero scores, and the majority were reading above 40 cwpm at endline.

Figure 20. Scatterplot of students' ORF scores at the initial assessment and endline (Nurturing Early Literacy—Maharashtra)



Overall, students in the Nurturing Early Reading project in Maharashtra demonstrated the highest reading proficiency across all project locations at both the initial assessment and endline. As the initial assessment was conducted toward the end of the project lifecycle, the treatment scores are likely a combination of the project's impact and other socioeconomic factors. Indeed, most students in Maharashtra fell into the highest socioeconomic group: 73% of students in control schools and 82% of students in treatment schools. Although the performance in reading comprehension was slightly more matched between the two groups compared to other project locations, at the initial assessment, students in the treatment group outperformed those in the control group by approximately 5 cwpm in the ORF subtask.

Table 22 displays the mean ORF and reading comprehension scores for students in treatment and control schools at the initial assessment and endline. It is important to note that the scores were not used to calculate the intervention impact. These values represent the mean scores recorded for the treatment and control groups at the initial assessment and endline without matched individual student scores.

The treatment group's average ORF scores at endline and the initial assessment were 45 cwpm and 35.3 cwpm, respectively. The mean ORF score for the control group at endline

was 43.7 cwpm, an increase from 30.2 cwpm at the initial assessment. Reading comprehension scores (% correct) improved for both the treatment and control groups; on average, students in the treatment group got 64% of questions correct, whereas those in the control group achieved 62% correct on average.

Table 22. Mean ORF and reading comprehension scores at the initial assessment and endline across treatment groups (Nurturing Early Literacy—Maharashtra)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	30.2 (0.9)	43.7 (1.1)
	Treatment	35.3 (1.1)	45 (1.5)
Reading comprehension (% correct)	Control	40.5% (1.3%)	62.2% (1.4%)
	Treatment	44% (2%)	63.5% (2.1%)

Table 23 describes the percentage of students scoring zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. In both the treatment and control groups, the percentage of students scoring zero on reading comprehension decreased. However, the reductions observed for these two groups did not differ significantly. In the treatment group, the percentage of students who scored zero on the ORF subtask at endline was higher than that in the control group.

Table 23. Percentage of children scoring zero at the initial assessment and endline across treatment groups (Nurturing Early Literacy—Maharashtra)

Percentage of Students Scoring Zero	Treatment Group	Initial Assessment (SE)	Endline (SE)
ORF	Control	5.2% (1)	4.3% (0.8)
	Treatment	1.6% (0.6)	3.5% (0.6)
Reading comprehension	Control	17.7% (1.8%)	8% (1.2%)
	Treatment	20.3% (2.1%)	7.6% (1%)

4.3 Teacher Innovations in Practice—Uttar Pradesh

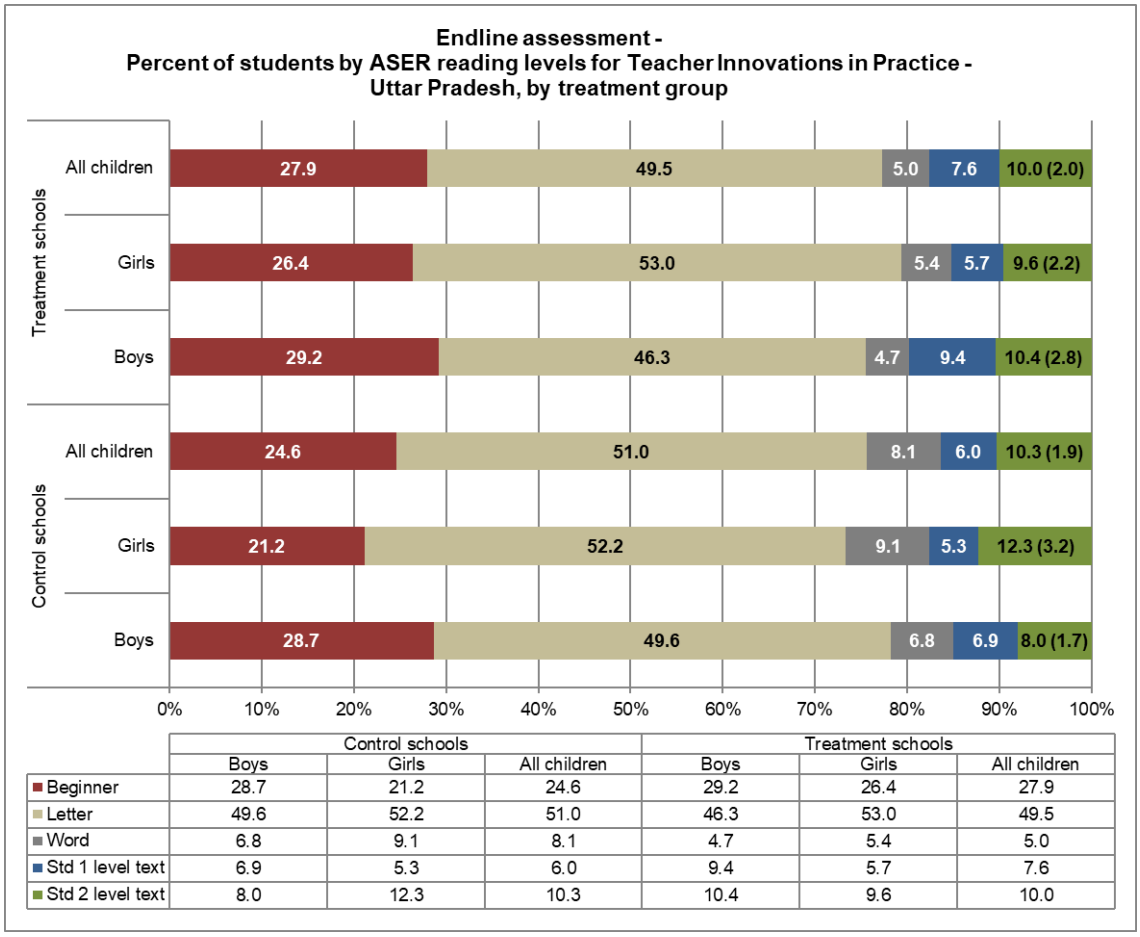
The Teacher Innovations in Practice project is implemented by STIR Education and is in the fourth and final year of implementation. The project began in Delhi and, with USAID support, expanded to Uttar Pradesh in October 2014. The 4-year project will conclude in September 2018. The project aims to reach 564,000 students in Standards 1–5. For the assessment, the project assessed students in Hindi from 10 districts in Uttar Pradesh.[7]

[7] Implementation roll out at the district level was as follows: one district in 2014 (Lucknow), three districts in 2015 (Raebareli, Varanasi, and Unnao), three districts in 2016 (Faizabad, Kanpur, and Jaunpur), and three districts in 2017 (Barabanki, Chandoli, and Mirzapur).

(1) ASER Results

Figure 21 shows the performance of Standard 2 students in treatment and control schools at the endline assessment for the Teacher Innovations in Practice project in Uttar Pradesh. Of the students in treatment schools, 27.9% could not even read letters at the endline assessment, compared to 24.6% of students in control schools. Approximately 10% of students in both treatment and control schools could read the Standard 2-level text.

Figure 21. Percentage of students by ASER reading level at endline (Teacher Innovations in Practice—Uttar Pradesh)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	28.7	21.2	24.6	29.2	26.4	27.9
Letter	49.6	52.2	51.0	46.3	53.0	49.5
Word	6.8	9.1	8.1	4.7	5.4	5.0
Std 1 level text	6.9	5.3	6.0	9.4	5.7	7.6
Std 2 level text	8.0	12.3	10.3	10.4	9.6	10.0

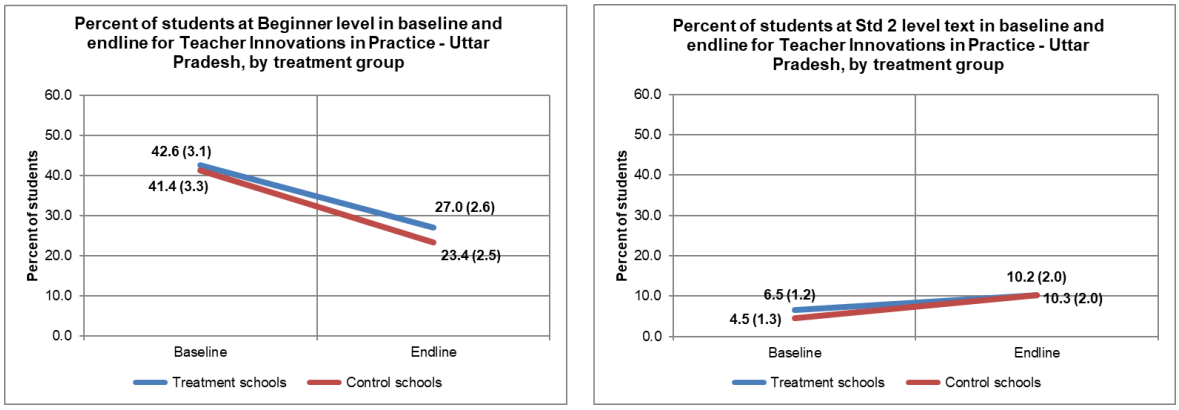
*SEs in parentheses

To understand the effect of the Teacher Innovations in Practice project in Uttar Pradesh on students' reading performance, we looked at the change in the proportion of students at the beginner and Standard 2 text levels from the initial assessment to the endline assessment for treatment and control schools.

From ***Figure 22***, we can see that the proportion of students at the beginner level decreased from 42.6% at the initial assessment to 27% at the endline assessment for treatment schools. For control schools, this proportion decreased from 41.4% to 23.4%. Regression analysis (see ***Annex H***) confirmed that the change in the proportion of students at the beginner level from the initial to endline assessments in treatment schools (15.6% points) did not differ significantly from that in control schools (18.0% points).

The proportion of students who could read the Standard 2-level text increased from 6.5% at the initial assessment to 10.2% at the endline assessment in treatment schools. In control schools, this proportion increased from 4.5% to 10.3%. The change in the proportion of students who could read the Standard 2-level text from the initial to endline assessments in treatment schools (3.7% points) did not differ significantly from that in control schools (5.9% points) (see ***Annex G***).

Figure 22. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (Teacher Innovations in Practice—Uttar Pradesh)



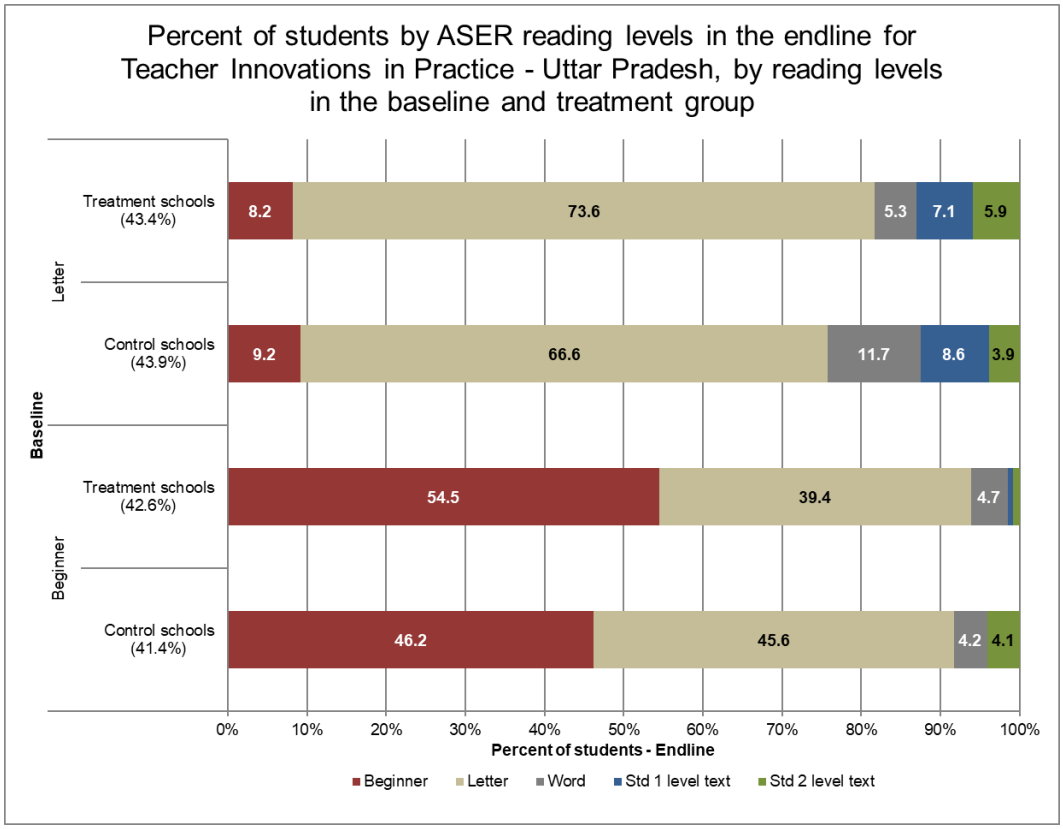
*SEs in parentheses

Another way to understand the effect of the Teacher Innovations in Practice project in Uttar Pradesh on reading performance is to see how well the program assisted students at lower levels of reading to progress to higher levels.

More than 40% of students in both the treatment (42.6%) and control schools (41.4%) were marked at the beginner level at the initial assessment. From ***Figure 23***, we can see that of the students who were marked at the beginner level at the initial assessment, 54.5% of those in treatment schools and 46.2% of those in control schools still could not even read letters. Another 45.6% of students in control schools who were marked at the beginner level at the initial assessment had progressed to the letter level, compared to 39.4% of such students in treatment schools.

At the initial assessment, more than 43% of students in both treatment (43.4%) and control schools (43.9%) were categorized at the letter level. Of these students, 73.6% of those in treatment schools remained at this level, and slightly more than 18% had progressed to the word level or higher at the endline assessment. In control schools, 66.6% of those students who were marked at the letter level at the initial assessment remained at the letter level, and more than 24% could read at the word level or higher at the endline assessment.

Figure 23. Change in the percentage of students at different ASER reading levels at endline (Teacher Innovations in Practice—Uttar Pradesh)



(2) EGRA Results

Table 24 displays the gain in mean ORF and reading comprehension scores and effect size for Standard 2 students by treatment group.

For Teacher Innovations in Practice—Uttar Pradesh, no significant gain in average student ORF or reading comprehension scores was found between the treatment groups. Both groups had gains in reading outcomes, but the difference in gains between treatment and control schools was not significant.

Table 24. Average student gains in ORF and reading comprehension scores (Teacher Innovations in Practice—Uttar Pradesh)

Mean Gain in Score	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	3.5 (0.7)	-0.2	-0.04
	Treatment	3.3 (0.6)		
Mean gain in reading comprehension (% correct)	Control	3.4% (0.8)	-0.4%	-0.04
	Treatment	3.0% (0.8)		

Figure 24 is a scatterplot of student scores at the initial assessment (x-axis) and endline (y-axis). A large proportion of students who scored zero remained scoring zero at endline across both treatment groups. While there were positive gains in means for both treatment groups, the scores remained low, reflecting poor reading outcomes. The majority of students are clustered at the lower, left-hand quadrant, with only a few achieving the benchmark at endline (***Section 5***). On average, students were reading at or below 7 cwpm.

Figure 24. Scatterplot of initial assessment and endline scores (Teacher Innovations in Practice—Uttar Pradesh)

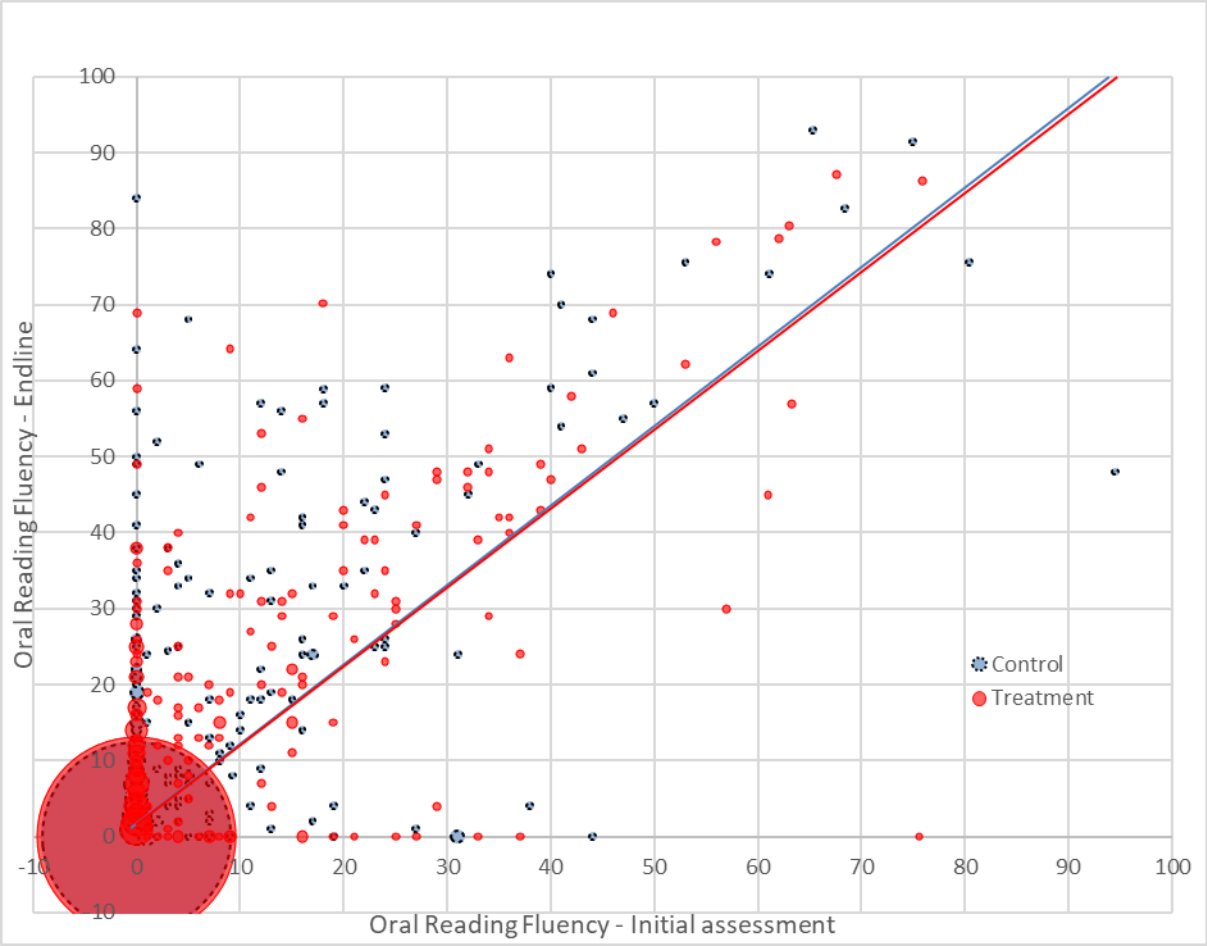


Table 25 displays the mean ORF and reading comprehension scores for the treatment and control groups at the initial assessment and endline. It is important to note that the scores were not used to calculate the intervention impact. These values represent the means of the treatment and control groups at the initial assessment and endline without matched individual student scores.

The mean ORF and reading comprehension scores improved for both the treatment and control groups from the initial to endline assessments. However, for both the treatment and control groups, the mean ORF was less than 10 cwpm, and on average, students scored approximately 6.5% on the reading comprehension subtask.

Table 25. Mean ORF and reading comprehension scores at the initial assessment and endline across treatment groups (Teacher Innovations in Practice—Uttar Pradesh)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	3.1 (0.6)	6.3 (1)
	Treatment	3.7 (0.6)	7.1 (1)
Reading comprehension (% correct)	Control	3.4% (0.8%)	6.5% (1.2%)
	Treatment	3.3% (0.6%)	6.4% (1%)

Table 26 describes the percentage of students who scored zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. In both groups, the percentages of students scoring zero on reading comprehension and ORF decreased. However, the observed reductions did not differ significantly between the treatment and control groups.

Table 26. Zero scores at the initial assessment and endline across treatment groups (Teacher Innovations in Practice—Uttar Pradesh)

Percentage of Students Scoring Zero	Treatment Group	Initial Assessment (SE)	Endline (SE)
ORF	Control	82.4% (2.5)	69.6% (2.9)
	Treatment	81.8% (2.7)	69.5% (3.1)
Reading comprehension	Control	92.6% (1.6%)	86.9% (2.3%)
	Treatment	92.2% (1.3%)	85.3% (2%)

4.4 Start Early: Read in Time

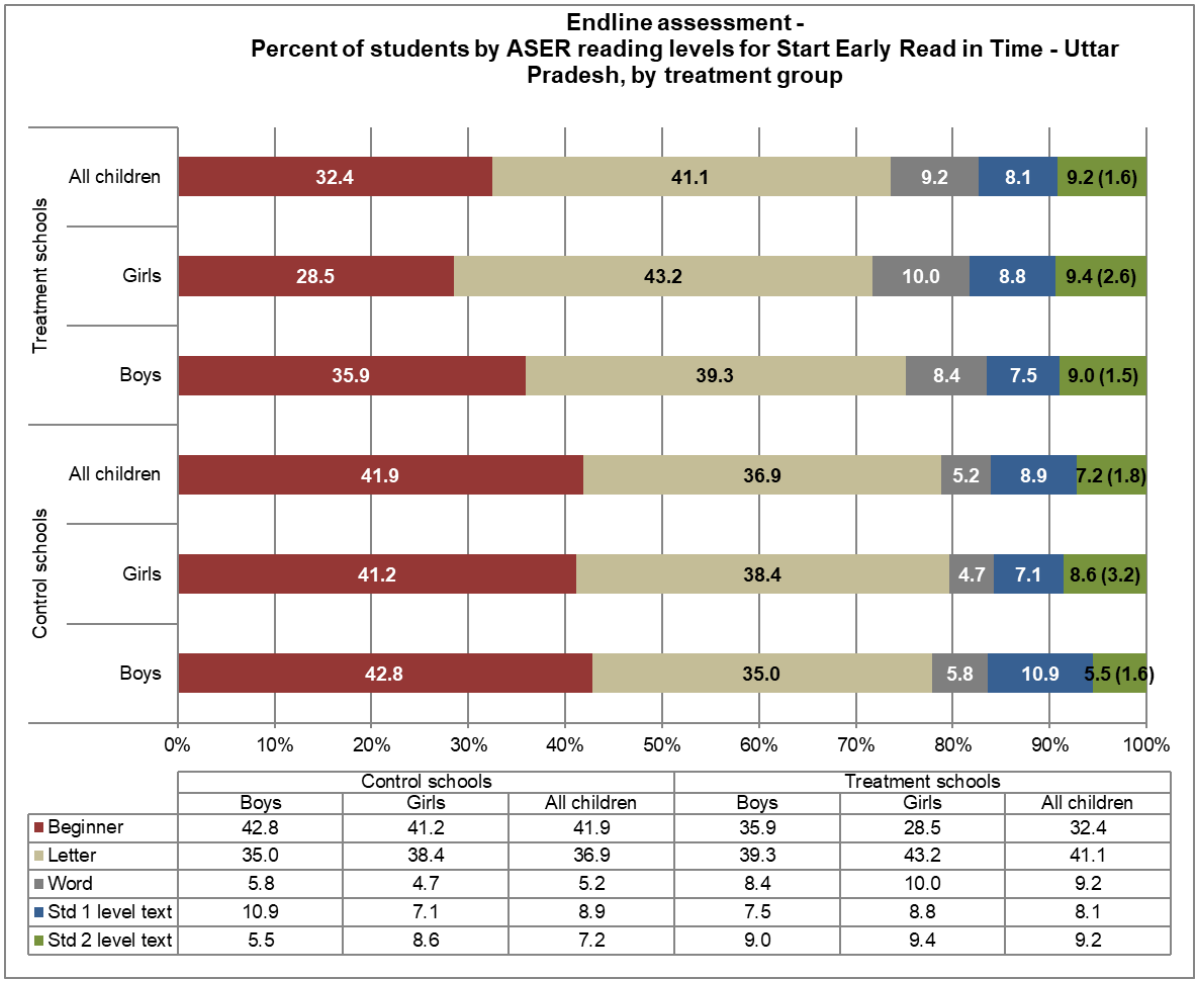
CARE India: India Solutions for Sustainable Development (CISSD) started implementing Start Early: Read in Time in July 2014. The project is currently in the fourth and final year of implementation and focuses on enhancing the reading skills of Standard 1–4 students in formal primary schools in Uttar Pradesh and Odisha. The aim is to reach 100,000 students among the most marginalized children (especially girls) in formal government primary schools. The reading assessment was administered in Hindi in five districts in Uttar Pradesh and in Oriya in one district in Odisha.

4.4.1 Start Early: Read in Time—Uttar Pradesh

(1) ASER Results

Figure 25 shows the performance of Standard 2 students in treatment and control schools at the endline assessment for the Start Early: Read in Time project in Uttar Pradesh. Of the students in treatment schools, 32.4% could not even read letters at the endline assessment, compared to 41.9% of students in control schools. Additionally, 7.2% of students in control schools could read the Standard 2-level text, compared to 9.2% of students in treatment schools.

Figure 25. Percentage of students by ASER reading level at endline (Start Early: Read in Time—Uttar Pradesh)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	42.8	41.2	41.9	35.9	28.5	32.4
Letter	35.0	38.4	36.9	39.3	43.2	41.1
Word	5.8	4.7	5.2	8.4	10.0	9.2
Std 1 level text	10.9	7.1	8.9	7.5	8.8	8.1
Std 2 level text	5.5	8.6	7.2	9.0	9.4	9.2

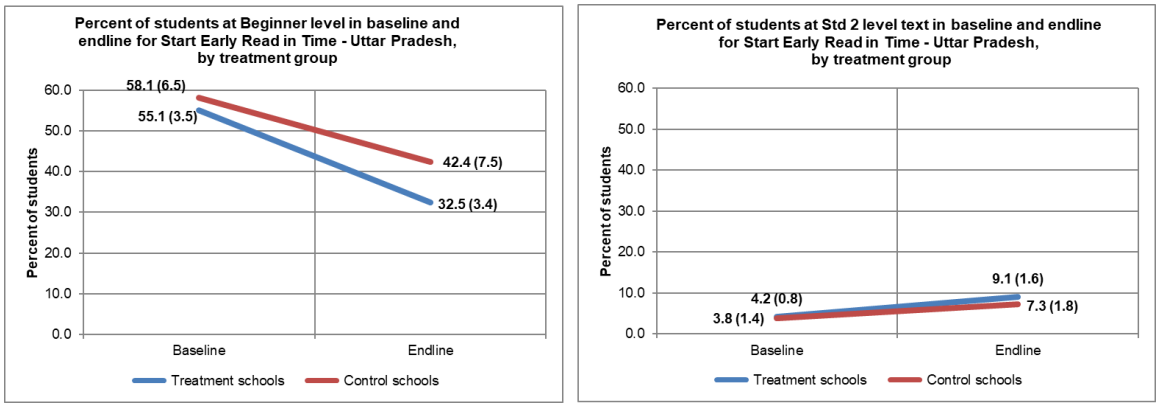
*SEs in parentheses

To understand the effect of the Start Early: Read in Time project in Uttar Pradesh on students' reading performance, we examined the change in the proportion of students at the beginner and Standard 2 text levels in treatment and control schools from the initial assessment to the endline assessment.

From ***Figure 26***, we can see that the proportion of students at the beginner level in treatment schools decreased from 55.1% at the initial assessment to 32.5% at the endline assessment. In control schools, this proportion decreased from 58.1% to 42.4%. Regression analysis (see ***Annex H***) confirmed that the change in the proportion of students at the beginner level from the initial to endline assessments did not differ significantly between treatment (22.6% points) and control schools (15.7% points).

Figure 26 shows that the proportion of students who could read the Standard 2-level text increased from 4.2% at the initial assessment to 9.1% at the endline assessment for treatment schools. This increase was from 3.8% to 7.3% in control schools. The change in the proportion of students who could read the Standard 2-level text from the initial to endline assessments did not differ significantly between treatment (4.9% points) and control schools (3.4% points) (see ***Annex G***).

Figure 26. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (Start Early: Read in Time—Uttar Pradesh)



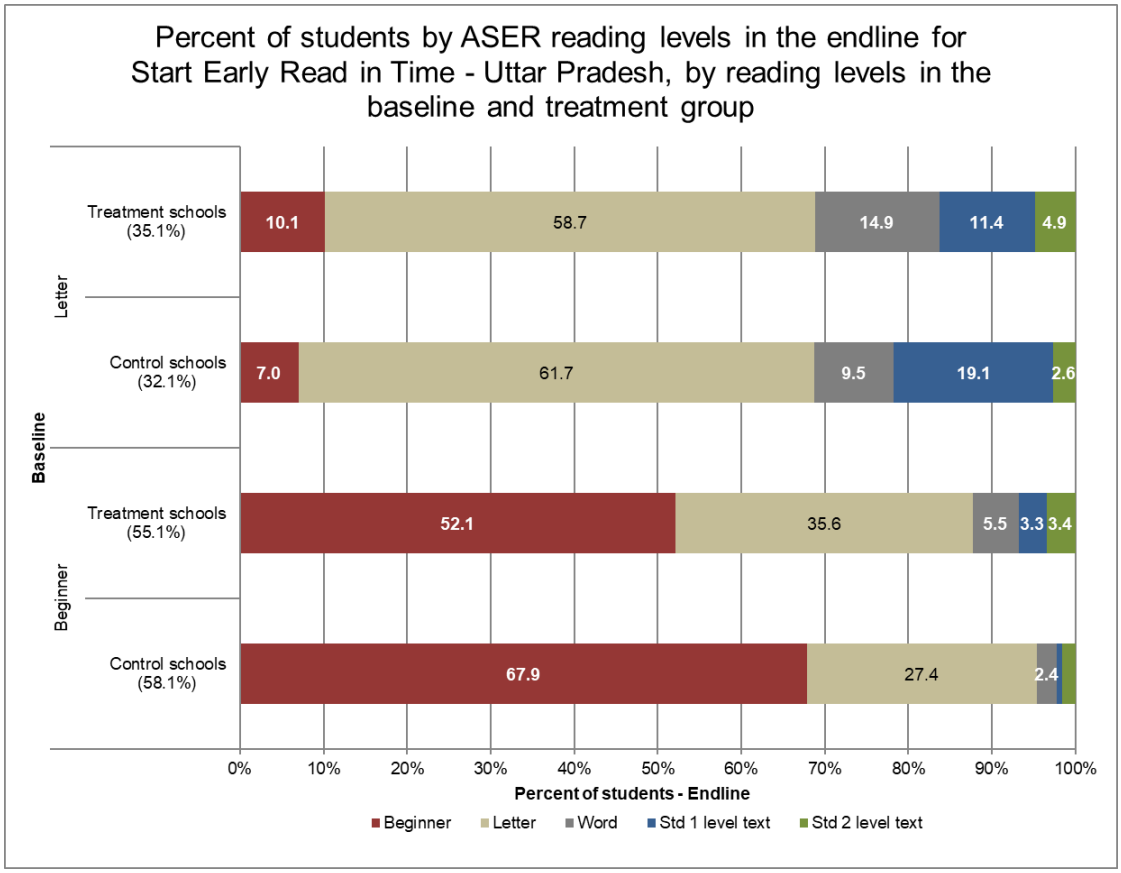
*SEs in parentheses

Another way to understand the effectiveness of the Start Early: Read in Time project in Uttar Pradesh on reading performance is to see how well the program assisted students at lower levels of reading to progress to higher levels.

At the initial assessment, 55.1% and 58.1% of students in treatment and control schools, respectively, could not even read letters. From ***Figure 27***, we can see that of the students who were marked at the beginner level, 67.9% remained at the beginner level in control schools, compared to 52.1% in treatment schools. In treatment schools, 35.6% of the students who were marked at the beginner level at the initial assessment had progressed to the letter level at the endline assessment, compared to 27.4% students in control schools.

In treatment and control schools, 35.1% and 32.1% of students, respectively, were marked at the letter level at the initial assessment. ***Figure 27*** shows that of these students, 58.7% of those in treatment schools remained at the letter level at the endline assessment, and 14.9% had progressed to the word level; these proportions are 61.7% and 9.5%, respectively, in control schools. In control schools, almost 22% of students marked at the letter level at the initial assessment were categorized at the Standard 1 text level or higher at endline, compared to slightly over 16% in treatment schools.

Figure 27. Change in the percentage of students at different ASER reading levels at endline (Start Early: Read in Time—Uttar Pradesh)



(2) EGRA Results

Table 27 displays the gains in mean ORF and reading comprehension scores and effect size for Standard 2 students by treatment group.

For Start Early: Read in Time—Uttar Pradesh, no significant gain in means was detected for either treatment group. Both groups had gains in reading outcomes, but the difference in gains between the treatment and control groups was not statistically significant.

For Start Early: Read in Time—Uttar Pradesh, the treatment group experienced gains over the control group in ORF (cwpm) and reading comprehension. The treatment group's mean scores increased on average by 3.7 cwpm between the initial assessment and endline, as compared to a mean gain of 2.8 cwpm in the control group. The impact of the intervention on the treatment group was that, on average, students were able to read nearly 1.0 cwpm more compared to the control. Similar results were seen for reading comprehension. The treatment group's mean scores increased on average by 3.3% between the initial assessment and endline, as compared to a mean gain of 1.2% in the control group. The impact of the intervention on the treatment group was that, on average, students gained 2% compared to the control. Similar to the gains found for ORF, the difference in the reading comprehension gain between the treatment and control groups was not statistically significant.

Table 27. Average student gains in ORF and reading comprehension scores (Start Early: Read in Time—Uttar Pradesh)

Mean Gain in Score	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	2.8 (0.8)	0.9	0.06
	Treatment	3.7 (0.7)		
Mean gain in reading comprehension (% correct)	Control	1.2% (1%)	2.1%	0.09
	Treatment	3.3% (0.8%)		

Figure 28 is a scatterplot of student scores at the initial assessment (x-axis) and endline (y-axis). A large proportion of students who scored zero remained scoring zero at endline in both treatment groups. Although both treatment groups saw positive gains in means, the scores remained low, reflecting poor reading outcomes. The majority of students are clustered at the lower, left-hand quadrant, with only a few achieving the benchmark at endline (see ***Section 5***); the average student was reading below 7 cwpm.

Figure 28. Scatterplot of initial assessment and endline scores (Start Early: Read in Time—Uttar Pradesh)

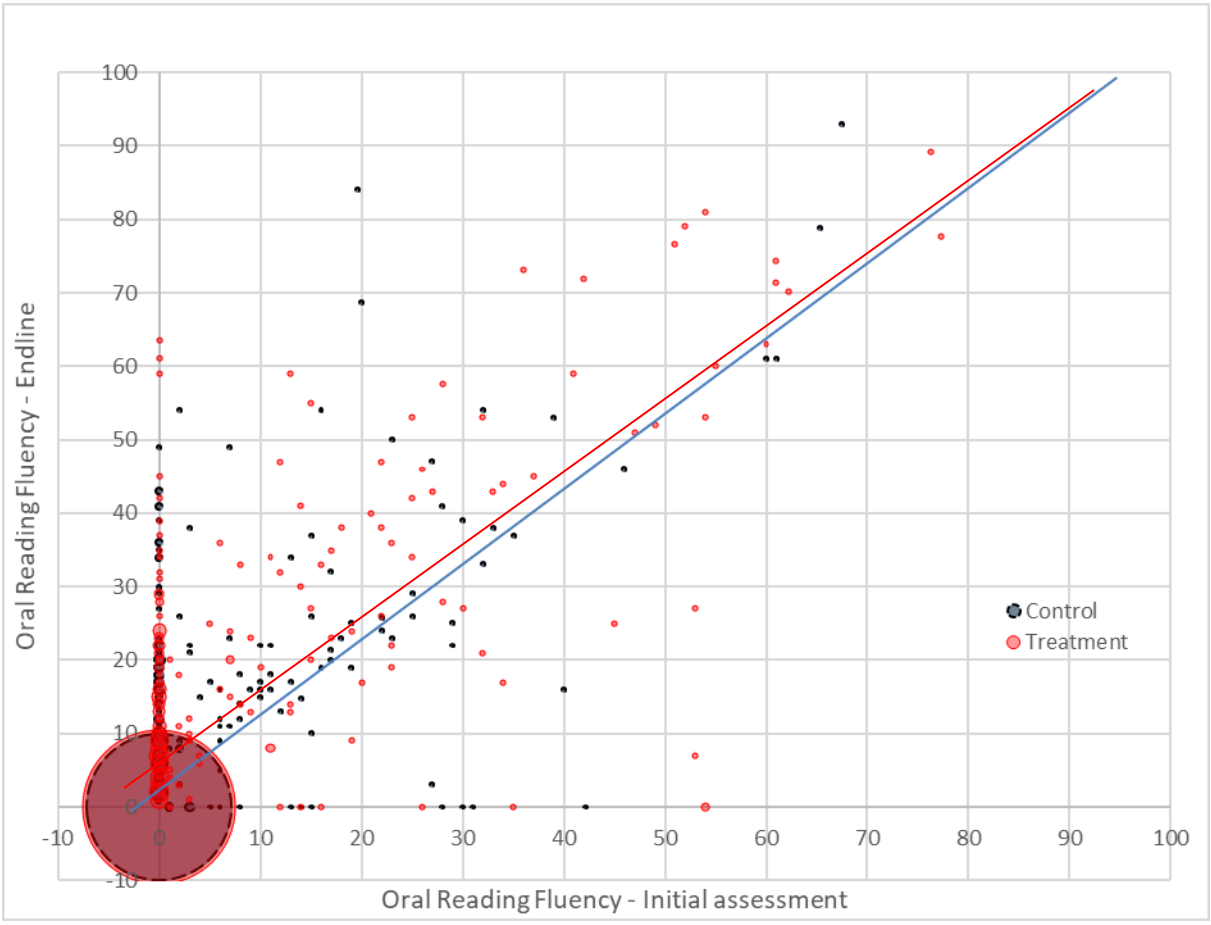


Table 28 displays the mean ORF and reading comprehension scores of students in the treatment and control groups at the initial assessment and endline. It is important to note that the scores were not used to calculate the intervention impact. These values represent the means of the treatment and control groups at the initial assessment and endline without matched individual student scores.

The average ORF and reading comprehension scores improved from the initial to endline assessments for both the treatment and control groups. The mean ORF scores for both treatment groups were approximately 2 cwpm at the initial assessment and improved to approximately 6 cwpm at endline. For reading comprehension, students in both treatment groups achieved less than 3% correct on average at the initial assessment, with these scores improving to 4% in the control group and 6% in the treatment group.

Table 28. Mean ORF and reading comprehension scores at the initial assessment and endline by treatment group (Start Early: Read in Time—Uttar Pradesh)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	2.4 (0.6)	5 (0.9)
	Treatment	2.4 (0.4)	6.6 (0.8)
Reading comprehension (% correct)	Control	2.5% (0.9%)	3.8% (0.9%)
	Treatment	2.8% (0.5%)	6.3% (0.9%)

Table 29 describes the percentage of students scoring zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. In both treatment groups, the percentage of students scoring zero had decreased at endline. However, the reductions observed did not differ significantly between the treatment and control groups.

Table 29. Percentage of children scoring zero at the initial assessment and endline by treatment group (Start Early: Read in Time—Uttar Pradesh)

Percentage of Students Scoring Zero	Treatment Group	Initial assessment (SE)	Endline (SE)
ORF	Control	89.2% (1.7)	76.4% (5.8)
	Treatment	89.3% (1.7)	69.3% (3.3)
Reading comprehension	Control	95.2% (1.3%)	89.4% (2.4%)
	Treatment	94.5% (1%)	85.6% (2.2%)

4.4.2 Start Early: Read in Time—Odisha

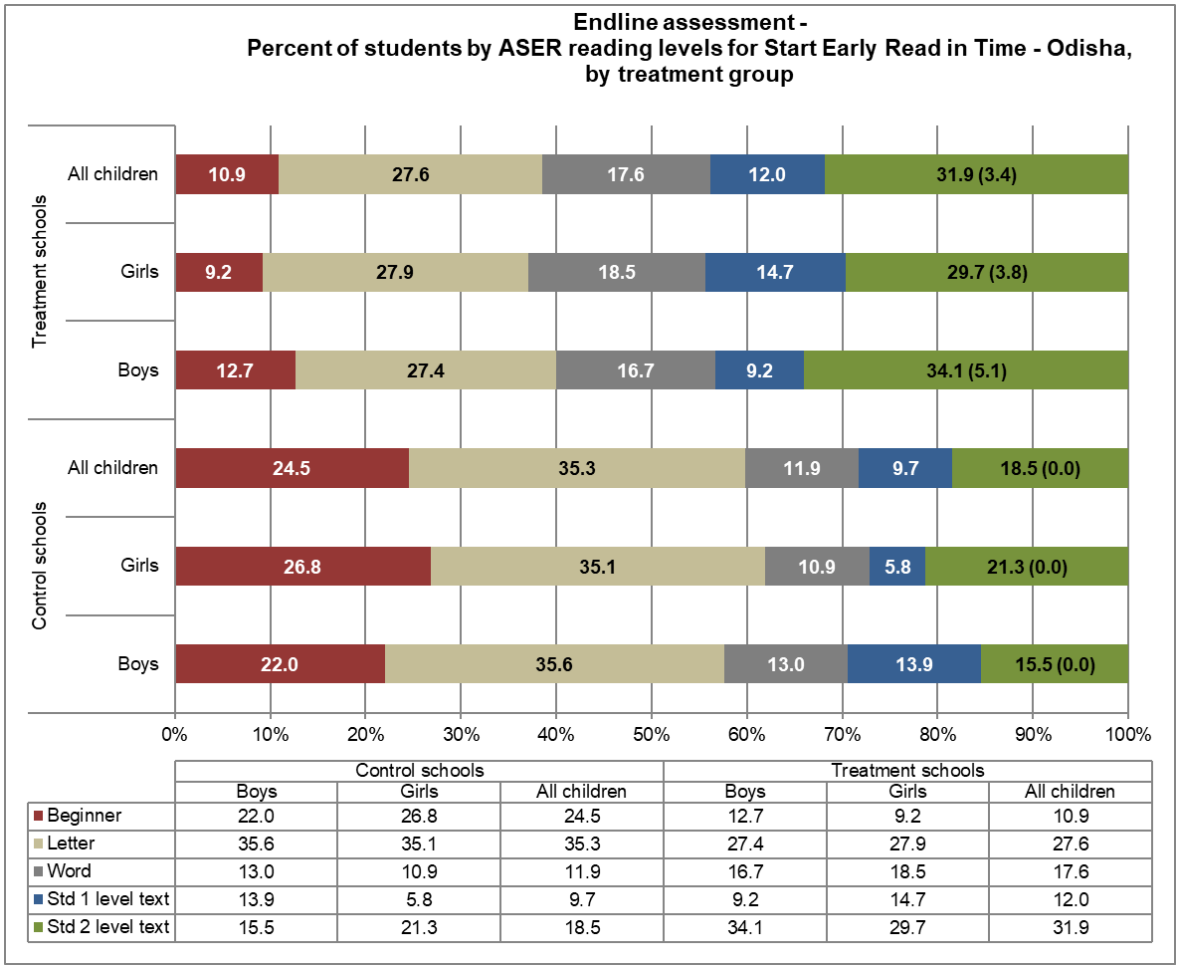
For the Start Early: Read in Time project in Odisha, 53% of the students assessed in the control group were girls, and 47% were boys. In treatment schools, equal numbers of boys and girls were assessed. The majority of students in both the control and treatment schools in Odisha reported speaking a language other than one of the six languages listed in the

assessment (63% and 52%, respectively). Only 30% of students in control schools and 45% in treatment schools reported speaking Oriya—the language of the assessment—at home.

(1) ASER Results

Figure 29 shows the performance of Standard 2 students in treatment and control schools at the endline assessment for the Start Early: Read in Time project in Odisha. Of the students in treatment schools, 31.9% could read the Standard 2-level text, and another 12% could read the Standard 1-level text but not the Standard 2-level text. In control schools, these proportions were 18.5% and 9.7%, respectively.

Figure 29. Percentage of students by ASER reading level at endline (Start Early: Read in Time—Odisha)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	22.0	26.8	24.5	12.7	9.2	10.9
Letter	35.6	35.1	35.3	27.4	27.9	27.6
Word	13.0	10.9	11.9	16.7	18.5	17.6
Std 1 level text	13.9	5.8	9.7	9.2	14.7	12.0
Std 2 level text	15.5	21.3	18.5	34.1	29.7	31.9

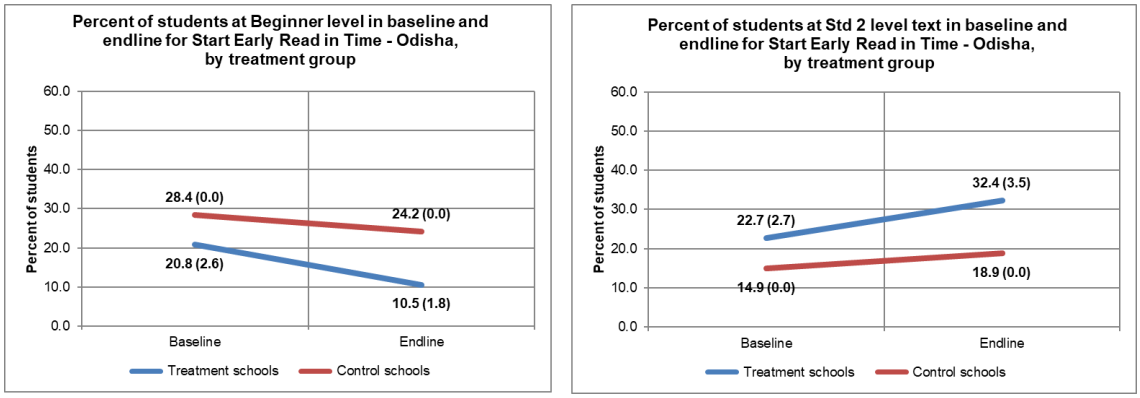
*SEs in parentheses

To understand the effect of the Start Early: Read in Time project in Odisha on students' reading performance, we examined the change in the proportion of students at the beginner and Standard 2 text levels in treatment and control schools from the initial assessment to the endline assessment.

From ***Figure 30***, we can see that the proportion of students at the beginner level decreased from 20.8% at the initial assessment to 10.5% at the endline assessment in treatment schools. In control schools, this proportion decreased from 28.4% to 24.2%. Regression analysis (see ***Annex H***) confirmed that the change in the proportion of students at the beginner level from the initial to endline assessments in treatment schools (10.3% points) did not differ significantly from that in control schools (4.3% points).

Figure 30 also shows that the proportion of students who could read the Standard 2-level text increased from 22.7% at the initial assessment to 32.4% at the endline assessment in treatment schools. The corresponding increase in control schools was from 14.9% to 18.9%. The change in the proportion of students who could read the Standard 2-level text from the initial to endline assessments in treatment schools (9.7% points) did not differ significantly from that in control schools (4.0% points) (see ***Annex G***).

Figure 30. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (Start Early: Read in Time—Odisha)



*SEs in parentheses

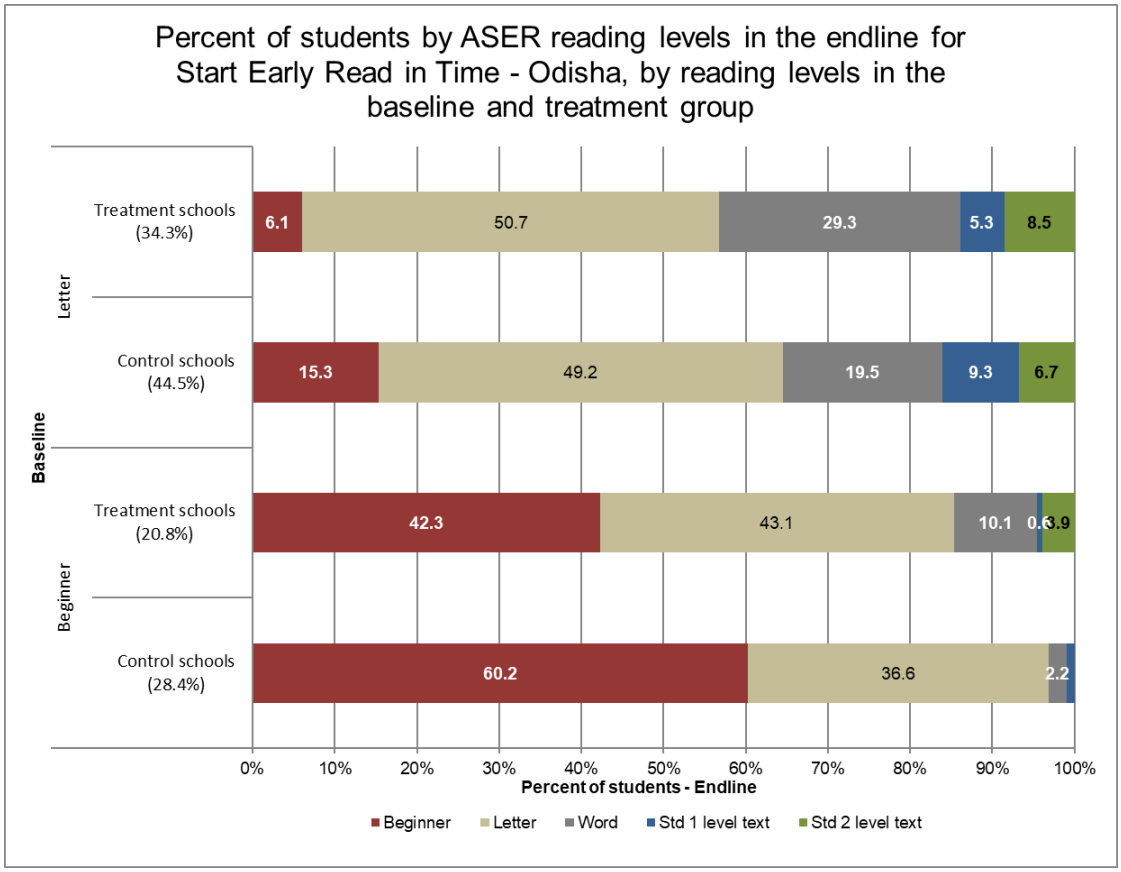
Another way to understand the effect of the Start Early: Read in Time project in Odisha on reading performance is to see how well the program assisted students at lower levels of reading to progress to higher levels.

At the initial assessment, 20.8% of Standard 2 students in treatment schools and 28.4% of students in control schools were marked at the beginner level. From ***Figure 31***, we can see that of these students, 60.2% of those in control schools were still at the beginner level, compared to 42.3% of those in treatment schools. In treatment schools, 43.1% and 10.1% of students who were marked at the beginner level at the initial assessment had progressed to the letter and word levels, respectively; in control schools, these proportions are 36.6% and 2.2%, respectively.

At the initial assessment, 34.3% of students in treatment schools were marked at the letter level, compared to 44.5% of students in control schools. Of these, 49.2% of students in control schools were still at the letter level at the endline assessment, compared to 50.7% of students in treatment schools. In treatment schools, 29.3% and approximately 14% of students who were categorized at the letter level at the initial assessment had progressed to the Word and Standard 1 text levels or higher, respectively. These proportions are 19.5% and 16% for control schools.

In control schools, 15.3% of letter-level students regressed to the beginner level from the initial assessment to the endline assessment. In comparison, in treatment schools, this proportion was 6.1%.

Figure 31. Change in the percentage of students at different ASER reading levels at endline (Start Early: Read in Time—Odisha)



(2) EGRA Results

For Start Early: Read in Time—Odisha, a positive impact attributable to the intervention was identified at endline: the treatment group saw greater gains in mean ORF scores between the initial assessment and endline. Treatment schools had a mean gain of 7.3 cwpm, whereas the control group had a mean gain of 3.4 cwpm; therefore, the treatment group's gain was, on average, 3.9 cwpm higher than that of the control group. This difference is significant at the 0.01 level, as indicated by **.

Table 30 displays the gains in mean ORF and reading comprehension scores and effect size for Standard 2 students in the treatment and control groups for Start Early: Read in Time—Odisha. At endline, students in the treatment group had a higher reading comprehension gain compared to the control group. Treatment schools had a mean gain of 7.2%, whereas the control group had a mean gain of 4.3%; therefore, the treatment group's gain was, on average, 2.9% higher than that of the control group. However, this difference was not found to be significant between the two treatment groups.

Table 30. Average student gains in ORF and reading comprehension scores (Start Early: Read in Time—Odisha)

Subtasks	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	3.4 (0.7)	3.9**	0.14
	Treatment	7.3 (0.9)		
Mean gain in reading comprehension (% correct)	Control	4.3% (1.4%)	2.9%	0.06
	Treatment	7.2% (1.1%)		

**Significant at the 0.01 level

Figure 32 plots student scores at the initial assessment (x-axis) and endline (y-axis). Compared to the other project location, Odisha has a wider spread of student scores, indicating that students represent a wide range of reading levels in treatment and control schools. Scores are not clustered near zero scores as they are in Odisha, and at endline, the average student read at 18.4 cwpm in the treatment group and 9.4 cwpm in the control group.

Regression lines were fitted for the treatment and control scatters separately. The treatment line (red) is above the control line (blue) and has a greater slope, confirming that, on average, students in treatment schools experienced greater individual gains.

Students in the treatment group had higher initial reading fluency and performed better than those in the control group at endline, as indicated by the clustering of red circles. The clustering of red dots on the y-axis indicates that the treatment was successful in shifting students away from zero scores.

The size of the blue circle at (0,0) indicates that a disproportionate number of students scored 0 cwpm at the initial assessment and remained there at endline (control). Notably, at the initial assessment, the mean ORF score for the control group was 5.8 cwpm (83% zero scores), compared to 10.8 cwpm for the treatment group (82% zero scores).

Figure 32. Scatterplot of initial assessment and endline scores (Start Early: Read in Time—Odisha)

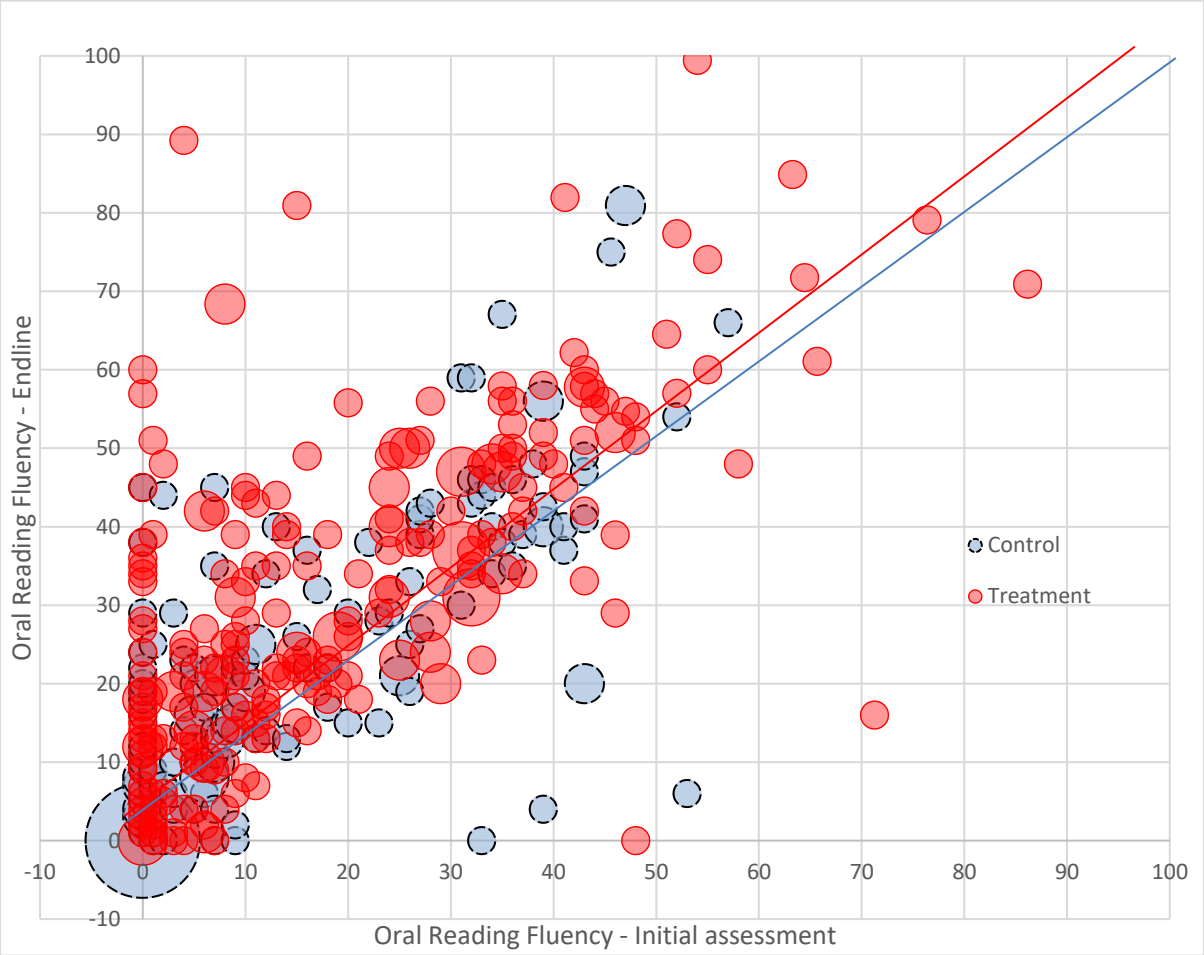


Table 31 displays the mean ORF and reading comprehension scores for the treatment and control groups at the initial assessment and endline. It is important to note that the scores were not used to calculate the intervention impact. The values represent the means of the treatment and control groups at the initial assessment and endline without matched individual student scores.

For both the treatment and control groups, the average ORF and reading comprehension scores improved from the initial to endline assessments. However, the improvement in average scores for the treatment group was greater on both measures. Whether the amount of improvement in either treatment or control group is adequate for the amount of elapsed instruction time between the initial and endline assessments cannot be determined.

Table 31. Mean ORF and reading comprehension scores at the initial assessment and endline across treatment groups (Start Early: Read in Time—Odisha)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	5.8 (0.8)	9.4 (1.2)
	Treatment	10.8 (1.4)	18.4 (1.8)
Reading comprehension (% correct)	Control	6.5% (0.9%)	10.8% (1.7%)
	Treatment	10.3% (1.6%)	17.7% (1.8%)

Table 32 describes the percentage of students who scored zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. For both the ORF and reading comprehension subtasks, significantly higher percentages of students scored zero in the control group than in the treatment group.

Table 32. Percentage of students scoring zero at the initial assessment and endline across treatment groups (Start Early: Read in Time—Odisha)

Percentage of Students Scoring Zero	Treatment Group	Initial Assessment (SE)	Endline (SE)
ORF	Control	62.9% (4%)	52.7% (4.1%)
	Treatment	48.7% (3.7%)	30.9% (3.2%)
Reading comprehension	Control	82.3% (2.2%)	74.2% (3.5%)
	Treatment	73.1% (3.6%)	59.5% (3.2%)

4.5 RightToRead

EnglishHelper's RightToRead project was implemented from September 2015 to September 2017 in Maharashtra, Gujarat, Delhi, Tamil Nadu, Telangana, Karnataka, and Punjab. The project reached more than one million students in Standards 1–8 across eight states in India. The 2-year project was extended by an additional year in 100 schools in Maharashtra and expanded to 300 new schools in West Bengal in 2017. Schools in only six districts were purposively sampled out of the 16 districts in which the program is being implemented in the state. These six districts represent more than 75% of the intervention schools. The initial data were collected in the first year of project implementation. Learners were assessed in English.

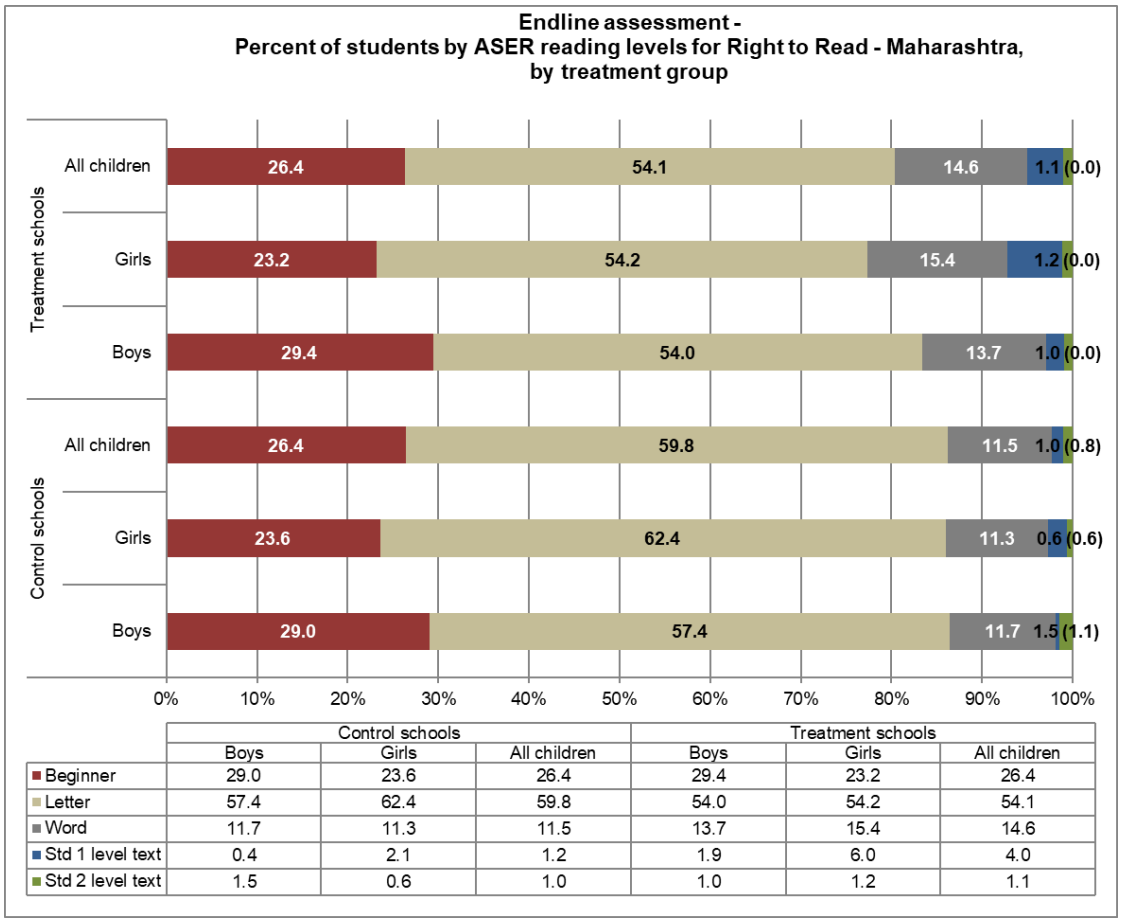
4.5.1 RightToRead—Maharashtra

(1) ASER Results

RightToRead's project in Maharashtra works to improve students' English abilities, and therefore, students were tested using an English reading tool for both the initial and endline assessments. ***Figure 33*** shows the performance of Standard 2 students in treatment and control schools at the endline assessment for this project.

At the endline assessment, 80.4% of students in treatment schools could not read at the word level or higher, compared to 86.3% of students in control schools. In both treatment and control schools, only approximately 1% of students could read the Standard 2-level text.

Figure 33. Percentage of students by ASER reading level at endline (RightToRead—Maharashtra)



	Control schools			Treatment schools		
	Boys	Girls	All children	Boys	Girls	All children
Beginner	29.0	23.6	26.4	29.4	23.2	26.4
Letter	57.4	62.4	59.8	54.0	54.2	54.1
Word	11.7	11.3	11.5	13.7	15.4	14.6
Std 1 level text	0.4	2.1	1.2	1.9	6.0	4.0
Std 2 level text	1.5	0.6	1.0	1.0	1.2	1.1

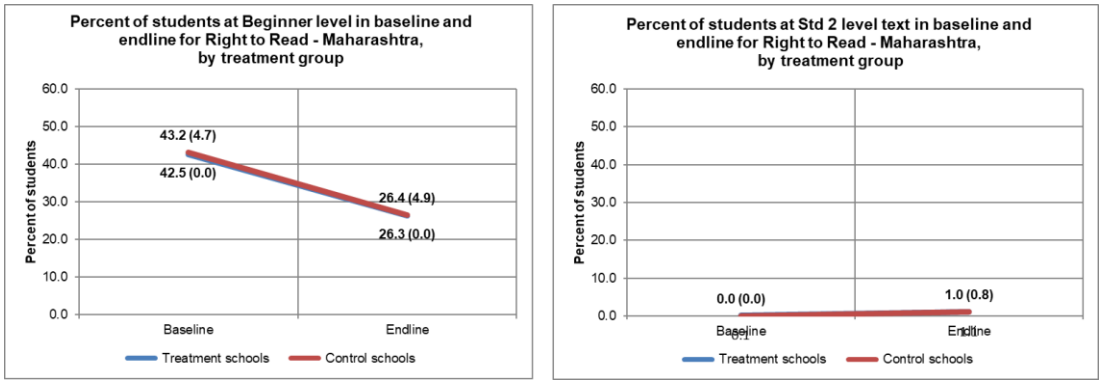
*SEs in parentheses

To understand the effect of the RightToRead project in Maharashtra on students' reading performance, we looked at the change in the proportion of students at the beginner and Standard 2 text levels in treatment and control schools from the initial assessment to the endline assessment.

From ***Figure 34***, we can see that the proportion of students at the beginner level in treatment schools decreased from 42.5% at the initial assessment to 26.3% at the endline assessment. In control schools, this proportion decreased from 43.2% to 26.4%. Regression analysis (see ***Annex H***) confirmed that the change in the proportion of students at the beginner level from the initial to endline assessments in treatment schools (16.2% points) did not differ significantly from that in control schools (16.8% points).

The proportion of students who could read the Standard 2-level text increased to approximately 1% in both treatment and control schools. The change in the proportion of students who could read the Standard 2-level text from the initial to endline assessments in treatment schools did not differ significantly from that in control schools (see ***Annex G***).

Figure 34. Percentage of students at the beginner and Standard 2 levels at the initial assessment and endline (RightToRead—Maharashtra)



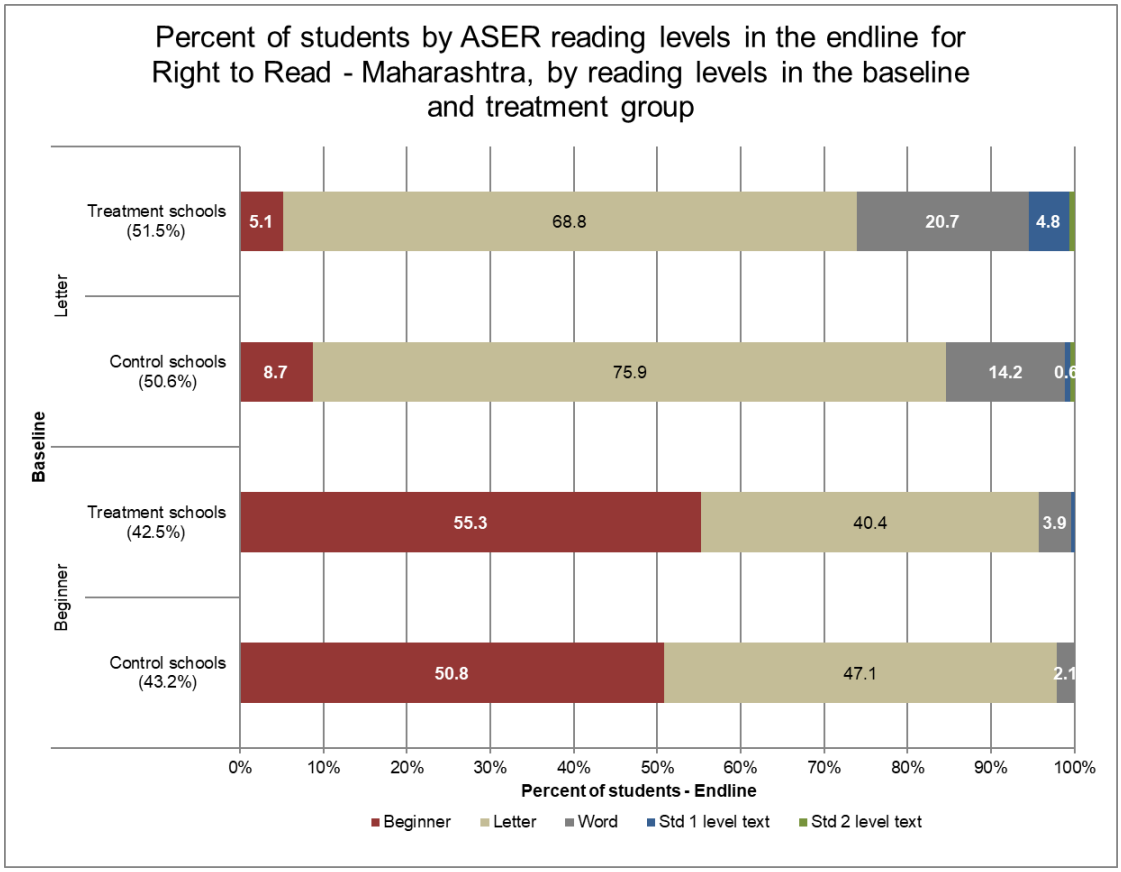
*SEs in parentheses

Another way to understand the effect of the RightToRead project in Maharashtra on reading performance is to see how well the program assisted students at lower levels of reading progress to higher levels.

At the initial assessment, more than 40% of students in both the treatment (42.5%) and control schools (43.2%) were marked at the beginner level. From ***Figure 35***, we can see that of these students, 50.8% of those in control schools remained at the beginner level, compared to 55.3% of those in treatment schools. In treatment schools, 40.4% of students who were marked at the beginner level at the initial assessment had progressed to the letter level, compared to 47.1% of students in control schools.

At the initial assessment, 51.5% of students in treatment schools and 50.6% of students in control schools were marked at the letter level. Of these students, 75.9% of those in control schools remained at the letter level at the endline assessment, compared to 68.8% of those in treatment schools. At the endline assessment, 20.7% and 14.2% of students who were categorized at the letter level in treatment and control schools, respectively, at the initial assessment had progressed to the word level. In control schools, 8.7% of letter-level students regressed to the beginner level from the initial assessment to the endline assessment. In comparison, this proportion was 5.1% in treatment schools.

Figure 35. Change in the percentage of students at different ASER reading levels at endline (RightToRead—Maharashtra)



(2) EGRA Results

For RightToRead—Maharashtra, a positive impact on mean ORF scores attributable to the intervention was identified at endline. The treatment group saw greater gains in mean ORF scores between the initial assessment and endline. Treatment schools had a mean gain of 3.3 cwpm, whereas the control group had a mean gain of 0.9 cwpm; therefore, the treatment group's gain was, on average, 2.4 cwpm higher than that of the control group. This difference is significant at the 0.01 level, as indicated by **.

Table 33 displays the gains in mean ORF and reading comprehension scores and effect size for Standard 2 students in the treatment and control groups. The control group outperformed the treatment group in terms of the mean gains in students' reading comprehension scores. On average, students improved very little, with no statistically discernable difference in the gains achieved in treatment and control schools.

Table 33. Average student gains in ORF and reading comprehension scores (RightToRead—Maharashtra)

Subtasks	Treatment Group	Average Student Gain (SE)	IMPACT Difference (T-C)	Effect Size
Mean gain in ORF (cwpm)	Control	0.9 (0.4)	2.4**	0.11
	Treatment	3.3 (0.5)		
Mean gain in reading comprehension (% correct)	Control	1.8% (0.9%)	-0.4%	-0.03
	Treatment	1.4% (0.4%)		

**Significant at the 0.01 level

Figure 36 is a scatterplot of student scores at the initial assessment (x-axis) and endline (y-axis). A large proportion of students who scored zero remained scoring zero at endline across both treatment groups. Although both groups saw positive gains in mean scores, the scores remained very low, reflecting poor reading outcomes. The majority of students are clustered at the lower, left-hand quadrant, with only a few achieving the benchmark at endline (see ***Section 5***); the average student was reading below 4 cwpm.

Figure 36. Scatterplot of initial assessment and endline scores (RightToRead—Maharashtra)

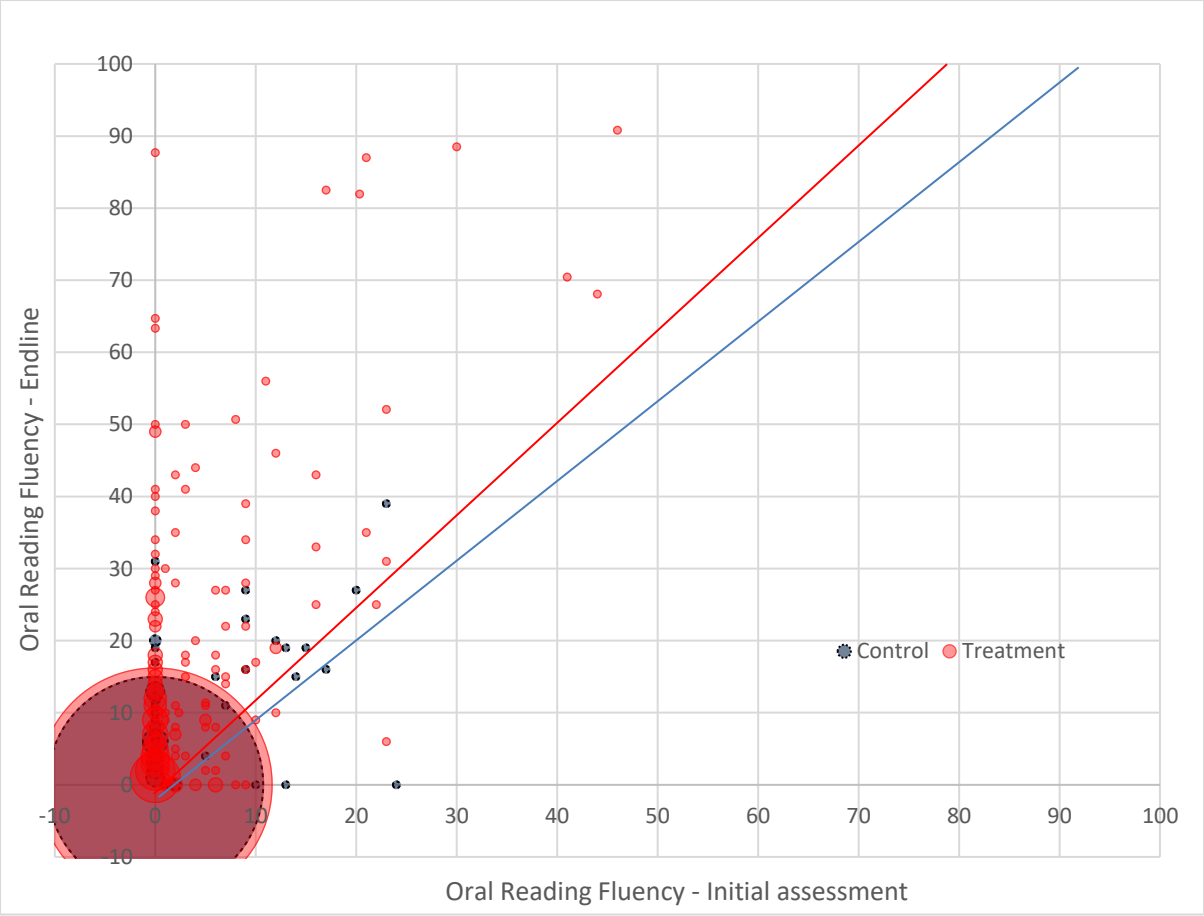


Table 34 displays the mean ORF and reading comprehension scores of the treatment and control groups at the initial assessment and endline. It is important to note that the scores were not used to calculate the intervention impact. These values represent the means of the treatment and control groups at the initial assessment and endline without matched individual student scores.

In both the treatment and control groups, the mean ORF scores measured for EnglishHelper—Maharashtra were the lowest recorded at the initial assessment. At endline, the reading comprehension scores improved to 2% for both treatment groups, although the majority of students were still unable to recognize a single word.

Table 34. Mean ORF and reading comprehension scores at the initial assessment and endline across treatment groups (RightToRead—Maharashtra)

Mean Scores	Treatment Group	Initial Assessment Mean (SE)	Endline Mean (SE)
ORF	Control	0.6 (0.2)	1.6 (0.7)
	Treatment	0.7 (0.1)	4 (0.6)
Reading comprehension (% correct)	Control	0.1% (0.1%)	1.9% (0.9%)
	Treatment	0.2% (0%)	1.6% (0.4%)

Table 35 describes the percentage of students who scored zero at the initial assessment and endline. Again, these scores are not matched student scores but straight averages for each treatment group at two time points. In both groups, the percentage of students scoring zero on reading comprehension and ORF decreased. However, the reduction in ORF zero scores was larger in the treatment group than in the control group. Given that students were tested in English, it is clear that the students were struggling in a language that was not their first; indeed, most students could not recognize a single word of English. However, the treatment group saw a significant reduction in zero scores between the initial assessment (93%) and endline (75%), indicating that the students were shifting away from being unable to read a single word.

The reductions observed in the proportions of reading comprehension zero scores were not significantly different between the treatment and control groups, which is to be expected in the context of such low ORF.

Table 35. Percentage of students scoring zero at the initial assessment and endline across treatment groups (RightToRead—Maharashtra)

Percentage of students scoring zero	Treatment Group	Initial assessment & SE	Endline & SE
ORF	Control	95% (1.7)	87.7% (4.2)
	Treatment	93.3% (1)	75% (2.6)
Reading comprehension	Control	99.6% (0.3%)	92.9% (2.7%)
	Treatment	99% (0.2%)	93.9% (1.2%)

5 Reading Benchmarks

To determine the impact of interventions on the percentage of children reading at or above reading benchmarks, the mean gain was calculated for each treatment group, as was the gain difference. Where the gain difference is positive, the treatment group outperformed the control group; asterisks indicate where gain differences were significant.

Final adopted benchmarks:

- Kannada—35 cwpm
- Hindi—35 cwpm
- Marathi—40 cwpm
- English—30 cwpm
- Oriya—30 cwpm

Table 36 presents the endline impact on the treatment group as measured by the gain in the percentage of students reaching the reading benchmark. For example, 25.6% more students reached the benchmark in the control group at endline, compared to 21.5% in the treatment group. Reading benchmarks were finalized for the projects during the Benchmarking Workshop held May 2–3, 2018, in New Delhi, India.[8] For two of the three CmF projects, the percentage of students who reached the reading benchmark was higher in the control group than in the treatment group; however, neither gain difference was statistically significant, indicating that there was no difference between the treatment and control groups in terms of the percentage of students who reached the benchmarks.

In four project locations Start Early: Read in Time—Odisha, Scaling Up Early Reading Intervention Project—Chhattisgarh, Scaling Up Early Reading Intervention Project—Uttarakhand and RightToRead—Maharashtra treatment outperformed control and gain differences were significant at the 0.01 level. The largest gain in the percentage of students reaching benchmark was for the Chhattisgarh project location with a gain difference in 18.1% followed by Uttarakhand with 10.7%.

Table 36. Percentage gains in the number of students reading at the benchmark level

Percentage of Students Reading at Benchmark	Treatment	Change in Percentage	IMPACT Difference T-C	Effect Size
Nurturing Early Literacy—Maharashtra				
Marathi 40-cwpm benchmark	Control	25.6%	-4.1%	0.02
	Treatment	21.5%		
Nurturing Early Literacy—Rajasthan				
Hindi 35-cwpm benchmark	Control	4.3%	-2.1%	0.03
	Treatment	2.2%		
Nurturing Early Literacy—Karnataka				
Kannada 35-cwpm benchmark	Control	1.1%	1.1%	0.02
	Treatment	2.2%		
Start Early: Read in Time—Odisha				
Oriya 30-cwpm benchmark	Control	3.0%	6.7%**	0.10
	Treatment	9.7%		
Start Early: Read in Time—Uttar Pradesh				
Hindi	Control	1.7%	1.3%	0.04

[8] For more information, please reference the Initial Assessment Report.

Percentage of Students Reading at Benchmark	Treatment	Change in Percentage	IMPACT Difference T-C	Effect Size
35-cwpm benchmark	Treatment	3.0%		
Scaling Up Early Learning Intervention—Chhattisgarh				
Hindi 35-cwpm benchmark	Control	6.9%	18.1%**	0.30
	Treatment	25.0%		
Scaling Up Early Learning Intervention—Uttarakhand				
Hindi 35-cwpm benchmark	Control	7.0%	10.7%**	0.15
	Treatment	17.7%		
RightToRead—Maharashtra				
English 30-cwpm benchmark	Control	0.3%	2.5%**	0.07
	Treatment	2.8%		
Teacher Innovations in Practice—Uttar Pradesh				
Hindi 35-cwpm benchmark	Control	4.5%	0.4%	0.04
	Treatment	4.1%		

Table 37 displays the benchmark data in a different way by presenting the percentage of students at or above the reading benchmarks at two different time points. It is important to note that these calculations are straight percentages at the initial assessment and endline for each treatment group. Unlike in the table above, students were not matched to calculate the longitudinal impact.

Table 37. Percentage of students reading at the benchmark at the initial assessment and endline

Region	Language	Treatment	Percentage of Students Reaching the Benchmark—Initial Assessment	Percentage of Students Reaching the Benchmark—Endline
Nurturing Early Literacy				
Rajasthan	Hindi	Control	0%	4%
		Treatment	0%	2%
Maharashtra	Marathi	Control	32%	56%
		Treatment	42%	60%
Karnataka	Kannada	Control	1%	2%
		Treatment	1%	3%
Start Early: Read in Time				
Uttar Pradesh	Hindi	Control	2%	4%
		Treatment	3%	6%
Odisha	Odiya	Control	9%	12%
		Treatment	17%	26%

Region	Language	Treatment	Percentage of Students Reaching the Benchmark—Initial Assessment	Percentage of Students Reaching the Benchmark—Endline
Scaling Up Early Learning Intervention				
Chhattisgarh	Hindi	Control	4%	11%
		Treatment	12%	37%
Uttarakhand	Hindi	Control	7%	14%
		Treatment	16%	34%
RightToRead				
Maharashtra	English	Control	0%	0%
		Treatment	0%	3%
Teacher Innovations in Practice				
Uttar Pradesh	Hindi	Control	3%	7%
		Treatment	4%	8%

6 Conclusion

As determined by the ASER analysis, three project locations are performing significantly better than control schools:

- Scaling Up Early Learning Intervention—Chhattisgarh,
- Scaling Up Early Learning Intervention—Uttarakhand, and
- Nurturing Early Literacy project in Karnataka.

To a large extent, the results of the EGRA impact analysis mimic those of the ASER analysis, which detected significant reading performance gains in both Scaling Up Early Learning Intervention project locations.

In the EGRA analysis, the treatment group outperformed the control group, and the gain differences between the treatment and control groups were significant at the 0.01 level for the following four project locations:

- Scaling Up Early Learning Intervention—Chhattisgarh,
- Scaling Up Early Learning Intervention—Uttarakhand,
- Start Early Read in Time—Odisha, and
- RightToRead—Maharashtra.

For the remaining project locations, although gains in reading outcomes may have been achieved, the differences in these gains between the treatment and control groups were not found to be statistically significant.

References

ASER Centre. (2016). *Annual Status of Education Report (Rural) 2016.* Retrieved December 25, 2017, from http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/aser_2016.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Ministry of Human Resource Department. (2017). *Department of School Education & Literacy, Government of India. Right to Education.* Retrieved December 25, 2017, from http://mhrd.gov.in/rte

SSA Shagun. (2017). *Department of School Education & Literacy, Ministry of Human Resource Development, Government of India. About SSA.* [Online]. Retrieved from http://ssashagun.nic.in/index.html

USAID/India. (2017). *Education* | *India*. Retrieved December 25, 2017, from https://www.usaid.gov/india/education

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014). *What Works Clearinghouse: Procedures and Standards Handbook (Version 3.0).* Retrieved from http://whatworks.ed.gov

Vagh, S. B. (2009). *Evaluating the reliability and validity of the ASER testing tools*. New Delhi, India: ASER Centre. Retrieved December 2, 2017, from http://www.asercentre.org/sampling/precision/reliability/validity/p/180.html

Vagh, S. B. (2016). Is simple, quick and cost-effective also valid? Evaluating the ASER Hindi Reading Assessment in India. In UNESCO Institute for Statistics (UIS). *Understanding what works in oral reading assessments: Recommendations from Donors, Implementers and Practitioners* (p. 202–212). Montreal: UNESCO Institute for Statistics. Retrieved December 25, 2017, from http://uis.unesco.org/sites/default/files/documents/understanding-what-works-in-oral-reading-assessments-2016-en.pdf

Annex A: Details of the Study Methodology

This annex describes the methodology used to collect, analyze, and weight the data from the endline assessment. The first section describes the study's sampling approach and procedures. The second section describes the weighting procedure.

A.1 Sampling Design and Procedures

To meet the objectives of this evaluation, a two-stage sampling design was used, with schools sampled in the first stage and students in the second stage. Because schools vary in size, schools were sampled for the initial assessment using probability proportional to size (PPS)[9] of Standard 2 enrollment. Student were sampled using simple random sample (SRS) of the Standard 2 enrollment register.

Sampling schools in the first stage: At least 60 treatment and 60 control schools were to be sampled per project location in the initial assessment. Treatment schools were sampled from the lists provided by the United States Agency for International Development (USAID)/India using the PPS sampling technique (on Standard 2 enrollment). When a program was spread over multiple districts within a state, a proportional sample was selected from each district. Ideally, control schools should be matched to the learning levels and other student characteristics, such as demographics and socioeconomic status (SES) in treatment schools. However, no such secondary data are available at the student level or even at the school level. The only information available at the school level is from the District Information System for Education (DISE). Given these constraints, the strategy used to select control schools in the initial assessment is provided below:

1. Sample treatment schools.
2. Determine the block (district sub-divisions) distribution of the sampled treatment schools.
3. Obtain the DISE list of all schools in the blocks of (2) above.
4. For each sampled treatment school, match a control school from the same block (within the same district) based on following criteria: school management type, standard present in school (primary only, Standards 1 to 5; primary with upper primary; Standards 6 to 8; and so on), enrollment in Standard 2, number of teachers appointed, and availability of computers and libraries for students' use.

Sampling of students in the second stage: In each sampled school (treatment and control), 20 students were sampled randomly from the Standard 2 enrollment register in the initial assessment. Of these students, 10 were boys and 10 girls. In the event a sampled child was not present in class on the day of the assessment, an additional child was sampled. In cases whereby class attendance was fewer than 10 by gender, all present students were sampled. However, because of either low enrollment and/or low attendance on the day of the school visit, the target sample of children was not met for many project locations.

A.1.2 Sampling of Additional Schools

Originally, 120 schools (60 treatment and 60 control) were to be sampled in each project location. However, during data collection for the initial assessment, a real-time online system

[9] PPS is a sampling technique in which the probability of selecting a sampling unit (school in our case) is proportional to the size of its population. The following steps were taken while sampling schools. First, the cumulative enrollment by school was calculated. Second, the total enrollment of the schools in the sampling frame was divided by the number of sampling units (schools) to get the sampling interval (SI). Third, a random number between 1 and the SI is chosen. This is referred to as the random start (RS). The RS denotes the site of the first school to be selected from the cumulated enrollment. Fourth, the following series of numbers is formed: RS, RS+SI, RS+2SI, RS+3SI..... The schools selected are those for which the cumulative enrollment contains the numbers in the series.

provided live updates regarding the number of students assessed in each project location. While monitoring these updates, the study team realized that for some project locations, field teams were not assessing enough students because of low enrollment and attendance in schools. Therefore, it was decided to sample additional schools in some project locations for the initial assessment.

Project locations where additional schools were sampled are provided below:

1. Scaling Up Early Reading Intervention—Uttarakhand: An additional 30 treatment and 30 control schools were sampled.
2. Nurturing Early Literacy—Maharashtra: An additional 10 treatment and 10 control schools were sampled.
3. Teacher Innovations in Practice—Uttar Pradesh: An additional 10 treatment and 10 control schools were sampled.
4. Start Early: Read in Time—Uttar Pradesh: An additional 10 treatment and 10 control schools were sampled.
5. Start Early: Read in Time—Odisha: An additional 10 treatment and 10 control schools were sampled.
6. RightToRead—Maharashtra: An additional 10 treatment and 10 control schools were sampled.

A.1.3 Incomplete Data Collection

Project locations where all sampled schools were not surveyed in the initial assessment are as follows:

1. Start Early: Read in Time—Odisha: Field teams were not allowed to carry out data collection in two (out of three) districts (Dhenkanal and Keonjhar).
2. RightToRead—Maharashtra: Permission was being arranged school by school and not as a blanket permission. Hence, many schools denied permission to collect data. As a result, 26 control schools in Pune and Latur districts were not surveyed.

For the endline assessment, the same schools were visited, and the same students were tracked and assessed again.

A.2 Weighting Procedures

The sample is meant to be representative for the program in a particular geography.

The two-stage design—PPS in the first stage and SRS in the second stage—yields a self-weighting sample at the program level. If, however, estimates are to be combined across programs and/or geographies, they would have to be weighted.

Sample design: Two stage:

1. Sample 60 schools from the treatment list using PPS on Standard 2 enrollment (from DISE).
2. Sample 20 Standard 2 students (10 boys and 10 girls) from the enrollment register using SRS.

Let,
N = Total number of Standard 2 students covered by the program;
S = Total number of schools in the program; and
n_s = Number of Standard 2 schools in school s.

So that,

$$\sum_{s=1}^{S} n_s = N$$

Then,

P_{ij} = Pr[student *i* in school *j* gets selected in the sample]
= Pr[school *j* gets selected] X Pr[student *i* gets selected]

$$= \frac{60\, n_j}{N} * \frac{20}{n_j} = \frac{1200}{N}$$

and,

w_{ij} = Weight associated with student *i* in school *j*
= inverse [P_{ij}]

$$= \frac{N}{1200}$$

Note that the weight associated with each student in a given program is the same—a self-weighting sample—and therefore, there is no need to weigh estimates at the program level. However, in practice, the procedure followed for sampling is often adjusted based on the ground reality. As a result, the weight calculations must take this into account, and it may be necessary to weigh estimates even at the district level. Some reasons why such weighting may become necessary are as follows:

1. The number of schools sampled in the first stage may vary across programs and districts. For instance, more than 60 schools may have to be sampled if schools have low enrollments. Where a program is spread over multiple districts, the sample is spread proportionately across the districts.
2. Although the second-stage sample is SRS from the enrollment register, in practice, students will be sampled from among those who are in attendance. Given the low attendance rates, it is quite possible that 20 students may not be sampled in each school.

As a result, the sample is no longer self-weighting, and estimates need to be weighted at every level.

Now,

P_{ijkp} = Pr[students *i* in school *j* of district *k* of program *p* get selected in the sample]
= Pr[school *j* of district *k* of program *p* gets selected] X Pr[student *i* gets selected]
= $n_{kp} \frac{e_j}{E_{kp}} \frac{n_{jkp}}{na_j}$

where,

n_{kp}= the number of schools sampled from district *k* of program *p*
e_j= the Standard 2 enrollment of school *j*
E_{kp}= the total Standard 2 enrollment of all schools of program *p* in district *k*
= $\sum_{m=1}^{N_{kp}} e_m$
N_{kp}= the total number of schools of program *p* in district *k*
n_{jkp}= the number of sampled students from school *j* of program *p* in district *k*

na_j= the number of students attending Standard 2 in school *j*

The weight associated with each student is simply the inverse of the probability of the student being selected in the sample. Or,

$w_{ijkp} = \frac{E_{kp}}{n_{kp}\, e_j}\ \frac{na_j}{n_{jkp}}$

Because the weight associated with a sampling unit represents the number of such units it represents in the population, the sum of the weights is the total population under consideration. Note that here, in the second stage, students are sampled from those who were attending school on the day of the survey during the initial assessment. Therefore, the sum of the weights of sampled students from district *k* of program *p* will be the total Standard 2 attendance of all the schools of program *p* in district *k*, not the total Standard 2 enrollment in these schools.

So,

$$\sum w_{ijkp} = \sum_{j=1}^{n_{kp}} \frac{E_{kp}}{n_{kp}\, e_j} \sum_{i=1}^{n_{jkp}} \frac{na_j}{n_{jkp}}$$
$$= \sum_{j=1}^{n_{kp}} \frac{E_{kp}}{n_{kp}\, e_j}\ na_j$$
$$= \frac{E_{kp}}{n_{kp}} \left[\frac{na_1}{e_1} + \frac{na_2}{e_2} + \cdots + \frac{na_{kp}}{e_{kp}}\right]$$
$$= E_{kp}\ \overline{(a/e)}$$

where $\overline{(a/e)}$ is the average attendance rate in the sampled schools of program *p* in district *k*.

A.2.1 Calculation of Weights in Control Schools

Assigning weights to the control schools is more problematic because these schools were not sampled. Instead, for each treatment school, a matching control school was identified from the DISE frame based on the criteria discussed above. There are two alternatives here:

1. The matched control group gets a “natural” weight and associated finite population correction (fpc) using the same methodology as the treatment group; this would assume that a two-stage sampling strategy was used to sample the control schools, even though there was no sampling in the first stage (second-stage sampling was identical to that performed in treatment schools).

2. The matched control “inherits” the survey weight of the treatment group to which it was matched.

It is not clear which method is better, and each has problems associated with the underlying assumptions. The sum of the weights, if we were to use the first procedure, would give us the population of Standard 2 students in the larger geography of the program area (i.e., the block/district from which the control schools were selected), and the control schools would have a much larger weight because they represent a larger universe. In the second case, the sum of the weights would approximate twice the population of the treatment area. In consultation with RTI, the first option was adopted.

A.2.2 Correction for Finite Population

This procedure is needed because both in the first stage (treatment) and the second stage, the population under consideration is small. The fpc simply the sampling rate at each stage of sampling and is invariant within the stratum (district). Therefore, the fpc for stage 1 in treatment schools is given by

$$fpc_1 = \frac{\#\ of\ sampled\ treatment\ schools}{Total\ \#\ of\ schools\ in\ program}$$

and similarly, for control schools, the fpc for stage 1 is given by

$$fpc_1 = \frac{\#\ of\ sampled\ control\ schools}{Total\ \#\ of\ schools\ in\ district - the\ schools\ in\ the\ program}$$

In the second stage, there is no difference between the treatment and control, and

$$fpc_2 = \frac{\#\ of\ sampled\ students}{Total\ \#\ of\ students\ attending\ school}$$

It is possible for fpc_2 to be greater than 1 if attendance is not recorded properly. In that case, it is set to 1.

A.2.3 Calculation of Endline Weights

Because this is a longitudinal sample, no sampling is performed for the endline assessment, and therefore, the weights and fpcs should remain unchanged. However, because of attrition, it may be necessary to recalculate the endline weights.

Two approaches can be discussed here. First, the endline can be treated as a fresh cross-section, and the cross-sectional weights can be calculated for this new wave; note that these weights would automatically be adjusted for attrition. However, these weights would sum to a different population—namely, average attendance during the endline assessment. In contrast, if the merged dataset were to be used, then which weight would be the appropriate weight is unclear. Therefore, the second approach would be to adjust the initial assessment weights for attrition. These adjusted weights would sum to the same population as the initial assessment, and with attrition, the weight associated with each tracked student would increase. Note that all other parameters used to set the survey design in Stata would remain unchanged at the initial assessment values.

Because attrition was excellent, creating endline cross-sectional weights by adjusting the initial assessment weights and using them for the gain scores provides enough diligence to the longitudinal design. In a merged initial assessment-endline dataset, we can use the new endline weights that account for the actual number of students at endline. The students without an endline assessment will not have the adjusted weights and will not be included in the estimates. Therefore, although the initial assessment and endline weights will be different, the population totals they represent should be approximately the same.

Annex B: Overview of United States Agency for International Development (USAID)/India Early Grade Reading (EGR) Project Descriptions

The information contained within this annex was provided by USAID/India on April 28, 2017.

1	**Start Early: Read in Time**	**CARE India: India Solutions for Sustainable Development (CISSD)**	**July 2014**	**July 2018**	**Uttar Pradesh, Odisha**
Description: The Start Early: Read in Time project aims to improve the reading skills of more than 100,000 students from marginalized communities attending government schools in the states of Uttar Pradesh and Odisha. The students, often first-generation students, come from a variety of cultural and linguistic backgrounds that can lead to a lack of support at school and at home. Data in both states indicate that after 5 years of schooling, only 44%–45% of students can read a Standard 2-level text. This project addresses these issues through systems strengthening, teacher training, and developing and disseminating teaching/learning materials that are contextualized to meet students' diverse needs. The key strategies are to build on students' prior linguistic knowledge and skills and to ensure a smooth transition from home language to school language. **Impact/Results:** In its first 2 years, the project has reached more than 100,000 students. Further, 5,000 more students can read with comprehension than when the project began. Examples of teaching and learning materials include poster stories, action cards, contextual story books written in mother tongue languages, and handbooks for teachers. This project is having an impact at the state level with widespread and state government endorsement of its position paper on early literacy and teacher support materials, state use of the project's model for convening teacher forums, and trainings for trainers at the state level. CISSD directly supports 480 schools in Uttar Pradesh and 516 schools in Odisha. Through state-level and systems strengthening activities, CISSD reaches 4,586 primary schools in Uttar Pradesh and 2,933 primary schools in Odisha. **Scalability:** This project is already having an impact at the state level, influencing more than 300,000 teachers working in government primary schools in Uttar Pradesh and Odisha.					
2	**Teacher Innovations in Practice**	**Schools and Teachers Innovating for Results (STIR)**	**October 2014**	**September 2018**	**Delhi, Uttar Pradesh**
Description: The Teacher Innovations in Practice project seeks to improve EGR outcomes in the states of Delhi and Uttar Pradesh by positively impacting the teaching practices of 14,657 teachers and the EGR achievements of 546,000 primary school students. The project motivates teachers by developing their mindsets, building an enabling environment, and enhancing their pedagogical skills and knowledge through micro-innovations, which lead to better student learning outcomes. Two examples of micro-innovations are creating class groups and awarding points for correct answers or playing reading games with flashcards. The program created a network of teachers (Teacher Changemaker Network) in which educators share and adopt successful micro-innovations and positively influence their peers to focus on changing classroom practices, resulting in improved student learning. After building strong Teacher Changemaker Networks, teachers are connected with relevant "next step" program partners, which can provide additional support. This program has leveraged funds from the Draper Richards Kaplan Foundation, Peery Foundation, Mulago Foundation, and Douglas Marshall Foundation and works closely with the state and local governments. **Impact/Results:** The program has already reached 10,038 teachers and 285,587 students across 5,156 primary schools. A total of 163 micro-innovations were identified, and STIR has launched 372 Teacher Changemaker Networks. **Scalability:** The Teacher Innovations in Practice project, implemented by STIR Education, began in Delhi and with USAID support has expanded to Uttar Pradesh. Now, it operates in 12 states across India.					

3	**RightToRead**	**EnglishHelper Education Technologies Private Limited**	**September 2015**	**September 2019**	**Maharashtra, West Bengal, Gujarat, Delhi, Tamil Nadu, Telangana, Karnataka, Punjab**
Description: Indian students from low income families study English within the constraints of first-generation literacy, regional disparities, and as an additional subject only (the first language in school is often not even their mother tongue) and, as such, are at a much lower level of English learning. This project addresses the shortage of English language teachers and poor English language skills among students in government-run primary schools. EnglishHelper uses an interactive computer program and digitized English language textbooks to help improve instruction and accessibility of the lessons. The software uses a combination of picture definitions, direct translations, and computer-generated narration to help with comprehension and pronunciation. The program has also leveraged funding from the Dell Foundation. **Impact/Results:** The program is reaching more than 1.1 million students across eight states. It is implemented in approximately 5,000 primary schools. **Scalability:** Since the launch of the project and subsequent visibility as a result of the USAID partnership, RightToRead now extends beyond the USAID partnership to four additional states in India and four other countries: Bangladesh, Sri Lanka, Sierra Leone, and Colombia. It was also endorsed by the State Institute of English, Government of Maharashtra, as an effective initiative that should be replicated in information and communication technology schools across the state.					
4	**Scaling Up Early Reading Intervention Project**	**Room to Read India Trust**	**September 2015**	**September 2020**	**Chhattisgarh, Uttarakhand**
Description: This program focuses on improving reading abilities for primary school students in the states of Chhattisgarh and Uttarakhand. According to the Annual Status of Education Report, only 25% of Standard 3 students can read a Standard 2-level text, emphasizing the need for reading skill development. The program improves how educators teach reading and instills good reading habits in the students. Examples include developing and disseminating teacher reference manuals with teaching instruction guidance and establishing libraries in schools. The project engages at the state level to ensure that the teaching and learning materials created align with the state government curriculum. It will also build systemic capacities and provide policy inputs that will enable the governments of Uttarakhand and Chhattisgarh to implement more-effective reading interventions. During the course of the project, 246,000 students will be reached directly. This project is expected to expand to include two other states in 2018 and indirectly benefit close to four million students in the long run. **Scalability:** The project engages at the state level to ensure that the teaching and learning materials created align with the state government curriculum. It also builds systemic capacities and provides policy inputs that will enable the governments of Uttarakhand and Chhattisgarh to implement the project model across all schools in the state. The project will expand to include two other states in 2018 and indirectly benefit close to four million students in the long run. With the right support, we are confident these programs can reach more students in more schools and improve their reading skills.					
5	**Nurturing Early Literacy Project**	**Centre for microFinance**	**October 2015**	**September 2019**	**Rajasthan, Maharashtra, Karnataka**
Description: The latest Annual Status of Education Report (ASER) released in 2015 shows that more than half of the government school students in Standard 5 are unable to read a Standard 2 text in their regional language. USAID partners with Tata Trusts and the Centre for microFinance to address this issue through implementing activities that build a strong foundation of emergent and early literacy competencies for 93,000 students across Rajasthan, Maharashtra, and Karnataka. The project aims to shift the prevalent rote-based pedagogy in India to one that views the child as an active student who can effectively learn sounds and symbols, read and write with comprehension, and apply their knowledge in everyday life. **Impact/Results:** Libraries in 100 schools in Rajasthan are now fully equipped, and there is an e-library pilot program in 10 schools. Teachers in 100 schools in Maharashtra have access to an online portal that provides them with resource materials and teaching modules they can use during class. Partners have also organized book fairs to engage the community and students.					

Annex C: Additional Attrition Information

For students who were not attending school on the day of endline assessment, data collectors made an attempt to track and assess them in their households. The tracking sheet also had location identifier variables related to students' households so that data collectors could identify students in the village communities if they were not attending school on the day of the visit. **Table C-1** below shows the overall and differential attrition between treatment and control schools for each project location.

Table C-1. Attrition by project location

Program	Assessed in Initial Assessment		Assessed in Initial and Endline Assessment		Attrition		Differential Attrition (Abs) [T - C]	Overall Attrition
	Student Sample in T Schools	Student Sample in C Schools	Student Sample in T Schools	Student Sample in C Schools	T schools	C schools		
Scaling Up Early Reading Intervention–Uttarakhand	974	707	881	657	9.5%	7.1%	2.5%	8.5%
Scaling Up Early Reading Intervention–Chhattisgarh	932	950	834	838	10.5%	11.8%	1.3%	11.2%
Nurturing Early Literacy–Rajasthan	666	591	573	532	14.0%	10.0%	4.0%	12.1%
Nurturing Early Literacy–Karnataka	1,039	783	933	714	10.2%	8.8%	1.4%	9.6%
Nurturing Early Literacy–Maharashtra	814	656	771	628	5.3%	4.3%	1.0%	4.8%
Teacher Innovations in Practice–Uttar Pradesh	896	869	782	764	12.7%	12.1%	0.6%	12.4%
Start Early Read in Time - Uttar Pradesh	946	826	848	729	10.4%	11.7%	1.4%	11.0%
Start Early Read in Time - Odisha	497	408	426	326	14.3%	20.1%	5.8%	16.9%
RightToRead–Maharashtra	1,064	752	968	682	9.0%	9.3%	0.3%	9.1%
Total	7,828	6,542	7,016	5,870	10.4%	10.3%	0.1%	10.3%

Table C-2 shows the number of students tracked for each project location by school type.

Table C-2. Number of students tracked in endline assessment by project location

Program	Students Surveyed for the Initial Assessment			Students Tracked for the Endline Assessment			
	Treatment Schools	Control schools	Total*	Treatment Schools	Control Schools	Total*	% Tracked
Scaling Up Early Reading Intervention—Uttarakhand	974	707	1,681	884	659	1,543	91.8
Scaling Up Early Reading Intervention—Chhattisgarh	933	957	1,890	836	846	1,682	89.0
Nurturing Early Literacy—Rajasthan	674	598	1,272	581	537	1,118	87.9
Nurturing Early Literacy—Karnataka	1,040	783	1,823	938	717	1,655	90.8
Nurturing Early Literacy—Maharashtra	815	660	1,475	775	632	1,407	95.4
Teacher Innovations in Practice—Uttar Pradesh	900	893	1,793	802	797	1,599	89.2
Start Early: Read in Time—Uttar Pradesh	951	839	1,790	864	752	1,616	90.3
Start Early: Read in Time—Odisha	498	419	917	431	344	775	84.5
RightToRead—Maharashtra	1,067	755	1,822	985	690	1,675	91.9
Total	**7,852**	**6,611**	**14,463**	**7,096**	**5,974**	**13,070**	**90.4**

*In the absence of consent from students or parents, even students who were tracked were not tested.

Table C-3 shows where students were tracked during the endline assessment.

Table C-3. Location of tracking of students by project location

Program	Treatment Schools				Control Schools			
	Students Tracked	Location of Tracking (% of Students)			Students Tracked	Location of Tracking (% of Students)		
		School	Home	Missing		School	Home	Missing
Scaling Up Early Reading Intervention—Uttarakhand	884	89.6	7.8	2.6	659	82.1	10.6	7.3
Scaling Up Early Reading Intervention—Chhattisgarh	836	90.4	9.3	0.2	846	89.5	8.6	1.9
Nurturing Early Literacy—Rajasthan	581	94.0	5.7	0.3	537	94.2	5.2	0.6
Nurturing Early Literacy—Karnataka	938	88.9	11.1	0.0	717	91.4	8.2	0.4
Nurturing Early Literacy—Maharashtra	775	98.6	1.3	0.1	632	98.6	0.2	1.3
Teacher Innovations in Practice—Uttar Pradesh	802	77.1	21.7	1.2	797	76.8	22.7	0.5
Start Early: Read in Time—Uttar Pradesh	864	86.1	12.8	1.0	752	85.4	14.4	0.3
Start Early: Read in Time—Odisha	431	83.5	14.6	1.9	344	80.8	18.3	0.9
RightToRead—Maharashtra	985	93.0	5.8	1.2	690	94.6	4.9	0.4
Total	**7,096**	**89.2**	**9.9**	**0.9**	**5,974**	**88.2**	**10.3**	**1.5**

Table C-4 shows the reason why students were not tracked for each project location.

Table C-4. Reasons for non-tracking of students by project location

Program	Treatment Schools						Control Schools					
	Students not Tracked	Reason for not Tracking (% of Students)					Students not Tracked	Reason for not Tracking (% of Students)				
		Not Present in Village	Home Closed	Left Village	Not Found	Other		Not Present in Village	Home Closed	Left Village	Not Found	Other
Scaling Up Early Reading Intervention—Uttarakhand	90	58.9	6.7	17.8	2.2	14.4	48	60.4	6.3	18.8	4.2	10.4
Up Early Reading Intervention—Chhattisgarh	97	47.4	8.2	41.2	1.0	2.1	111	40.5	9.0	50.5	0.0	0.0
Nurturing Early Literacy—Rajasthan	93	33.3	34.4	8.6	6.5	17.2	61	49.2	18.0	8.2	3.3	21.3
Nurturing Early Literacy—Karnataka	102	51.0	15.7	23.5	7.8	2.0	66	51.5	13.6	21.2	3.0	10.6
Nurturing Early Literacy—Maharashtra	40	67.5	10.0	5.0	2.5	15.0	28	28.6	17.9	25.0	0.0	28.6
Teacher Innovations in Practice—Uttar Pradesh	98	54.1	18.4	14.3	7.1	6.1	96	55.2	21.9	8.3	6.3	8.3
Start Early: Read in Time—Uttar Pradesh	87	72.4	13.8	4.6	4.6	4.6	87	63.2	18.4	5.7	8.0	4.6
Start Early: Read in Time—Odisha	67	47.8	29.9	4.5	14.9	3.0	75	44.0	40.0	2.7	12.0	1.3

Program	Treatment Schools						Control Schools					
	Students not Tracked	Reason for not Tracking (% of Students)					Students not Tracked	Reason for not Tracking (% of Students)				
		Not Present in Village	Home Closed	Left Village	Not Found	Other		Not Present in Village	Home Closed	Left Village	Not Found	Other
RightToRead—Maharashtra	82	54.9	22.0	9.8	4.9	8.5	65	32.3	32.3	7.7	20.0	7.7
Total	**756**	**53.2**	**17.7**	**15.7**	**5.7**	**7.7**	**637**	**48.4**	**19.8**	**17.4**	**6.4**	**8.0**

Annex D: Endline Data Collector Training and Data Collection Details

Table D-1 provides details on the content covered during the 3-day master trainer and state-level refresher data collector trainings. The refresher training was conducted prior to endline data collection to reacquaint data collectors with the reading assessments and administration procedures and protocols.

At the state level, approximately 400 participants were trained across various locations. More data collectors were trained than were needed, and based on performance and observations, participants were categorized as data collectors or monitors. A number of participants were dropped from consideration and were not selected for the actual data collection based on their performance during the training. Overall, 299 participants were selected as data collectors to participate in the fieldwork. An additional 87 participants served as data quality monitors.

Table D-1. Master trainer and state-level data collector training details

Day 1	Day 2	Day 3
About the United States Agency for International Development (USAID) and its funded projects	Pilot field practice: Practicing all the processes learned on Day 1	Quiz: Oral reading fluency (ORF) audio quiz to measure participant accuracy Quiz: Revision and clarifications
Introduction to the evaluation	Feedback session	Monitoring and recheck
Overview of the assessment process: • Preparation before going to school • Collecting school information • How to sample students in schools (initial assessment method) • Tracking and testing students who were assessed in the initial assessment using the tracking sheet • Completing the child information, child consent forms, and questionnaire • Annual Status of Education Report (ASER) testing • ORF reading and comprehension testing • Demonstration and discussion on the testing process	Quiz: Assessment process and ASER	State-level training and assessment planning

School-based practice. Master trainers and data collectors practiced collecting data by participating in a practice school visit during the training workshop. This practical experience allowed them to get hands-on experience with the actual data collection process and ensured logistics were well coordinated.

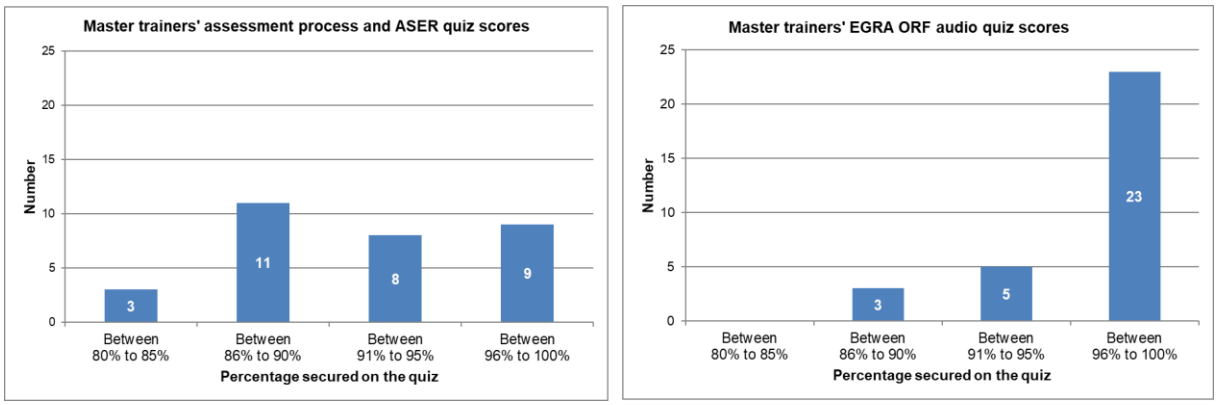
Master trainer and data collector performance. Central- and state-level trainings included an assessor accuracy measure to ensure that different data collectors had the ability to score the same student accurately on the ORF subtask, thus increasing the rigor and reliability of the measurement. Participants listened to an audio recording of a student reading the passage and answering the comprehension questions. Data from the assessor accuracy measure study were used to provide feedback to trainees, refining training where necessary and identifying trainees who did not reach a minimum performance standard. During the training, participants were also assessed on ASER protocols and the overall process of the assessment through a paper-based quiz. In trainings for the endline assessment, questions related to tracking of students were included in this quiz.

All master trainers scored above 80% on the assessment process and ASER quiz and over 85% on the ORF audio quiz.

Participants from the state-level trainings scoring less than 70% on the assessment process and ASER quiz and less than 80% on the ORF audio quiz were dropped. Additional clarification and training sessions were organized to fill the learning gaps identified through the quiz results. Furthermore, strict supervision of the lowest-performing data collectors was carried out in the initial days of the fieldwork.

Figure D-2 presents master trainers scores on the assessment process and ASER quiz and the ORF audio quiz during the endline refresher training.

Figure D-2. Master trainer scores on the assessment process and ASER quiz and the ORF audio quiz



Figures D-3 and ***D-4*** present data collectors' scores by project and location on the assessment process and ASER quiz and the ORF audio quiz, respectively.

Figure D-3. Data collector assessment process and ASER quiz scores

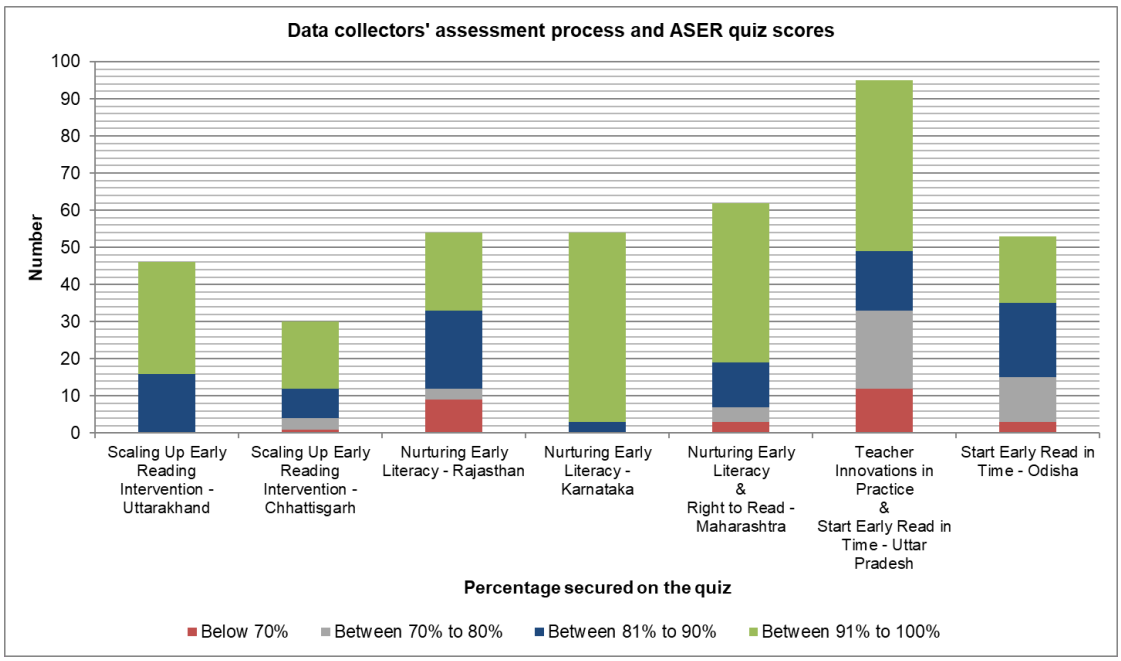


Figure D-4. Data collector ORF audio quiz scores

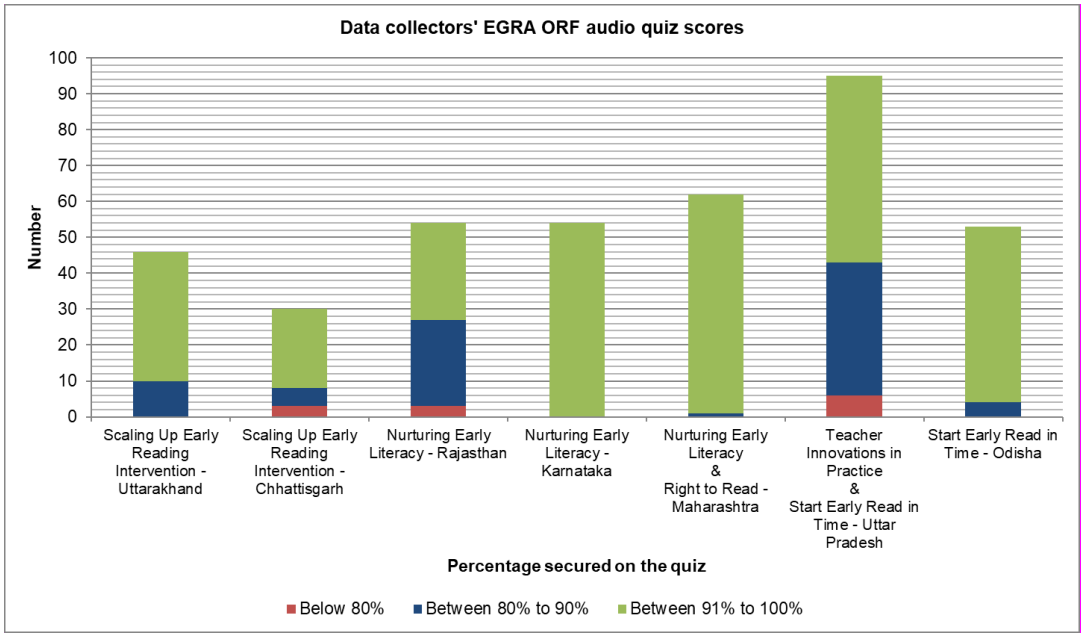


Table D-5 presents the number of data collectors and monitors selected by state.

Table D-5. Number of data collectors and monitors selected, by state

State	Participants on Day 1	Selected as Data Collectors	Selected as Monitors	Not Selected
Scaling Up Early Reading Intervention—Uttarakhand	46	34	12	0
Scaling Up Early Reading Intervention—Chhattisgarh	30	22	7	1
Nurturing Early Literacy—Rajasthan	57	38	9	10
Nurturing Early Literacy—Karnataka	56	42	12	2
Nurturing Early Literacy & RightToRead—Maharashtra	69	53	14	2
Teacher Innovations in Practice & Start Early: Read in Time—Uttar Pradesh	97	66	24	7
Start Early: Read in Time—Odisha	55	44	9	2
Total	**410**	**299**	**87**	**24**

Each field team comprised of the following people: Data Collector, Monitor, and Supervisor/ASER State Team. Approximately 25% of the participants in every project location were given the role of “Monitor” based on their performance during the training. The monitors visited a different school with the respective data collector each day and aided in ensuring data quality. Supervisors were mainly people such as ASER state team members and Pratham’s local staff in respective states/districts, individuals belonging to external organizations, and those who had helped in executing ASER surveys previously. These supervisors were responsible for the successful execution of the assessment in the states, including all quality control processes (e.g., monitoring, conducting desk and field checks).

For each project location, all the schools that were visited in the initial assessment were revisited at endline. In each of these schools, one data collector carried out the assessments. After reaching the school, the data collector met the Head Master and secured permission to conduct the assessment in the school. Data collectors were asked to implement simple activities with students in the class, such as introducing themselves to the students or playing a quick game before tracking or beginning the assessment. This was done to ensure that the students felt comfortable with the data collectors during the assessment. Next, the data collectors obtained information on enrollment and attendance for Standard 2. Data collectors tried to locate each student who was tested in the initial assessment using a tracking sheet with student details. For students who could not be tracked and tested in school, data collectors visited households.

Data collection timelines for each project location varied depending on the end-of-year exam calendar for the state, whether permission to carry out the survey in schools was received, and local holidays caused by festivals, among other factors. ***Table D-6*** shows the dates for data collector training and fieldwork by each project location.

Table D-6. Training and data collection timelines by project location

	Initial Assessment		Endline	
Program	**Training Dates**	**Data Collection Dates**	**Training Dates**	**Data Collection Dates**
Scaling Up Early Learning Intervention—Uttarakhand	September 13–15, 2017	September 16–October 13, 2017	February 11–13, 2018	February 14–24, 2018
Scaling Up Early Learning Intervention—Chhattisgarh	September 14–16, 2017	September 18–25, 2017	February 12–14, 2018	February 15–24, 2018
Nurturing Early Literacy—Rajasthan	September 13–15, 2017	September 16–25, 2017	March 20–22, 2018	March 23–31, 2018
Teacher Innovations in Practice—Uttar Pradesh	September 13–15, 2017	September 16–October 7, 2017	March 5–7, 2018	March 9–24, 2018
Start Early: Read in Time—Uttar Pradesh	September 13–15, 2017	September 16–October 9, 2017	February 13–15, 2018	February 16–24, 2018
Start Early: Read in Time–Odisha	September 15–17, 2017	September 17–25, 2017	March 20–22, 2018	March 23–31, 2018
Nurturing Early Literacy—Karnataka	September 14–17, 2017	September 19–23, 2017	February 12–14, 2018	February 15–24, 2018
Nurturing Early Literacy—Maharashtra	September 13–15, 2017	September 16–October 14, 2017	March 19–21, 2018	March 21–31, 2018
RightToRead—Maharashtra	September 13–15, 2017	September 16–October 14, 2017	February 11–13, 2018	February 14–24, 2018
School Excellence Program—Gujarat[10]	September 11–13, 2017	Not collected	Not collected	Not collected
RightToRead—West Bengal[11]	January 2018	Not collected	Not collected	Not collected

[10] School Excellence Program—Gujarat: State-level training for this project location was successfully completed on September 11–13, 2017, in Surat, Gujarat; however, data collection was not conducted for this project because the local government did not provide permission to access schools.

[11] Data collection was not conducted for this project because the local government did not provide permission to access schools.

Annex E: Quality Control/School Monitoring and Recheck

Monitoring and recheck provide a consolidated source of quality assurance for a project's progress. This process also supports evidence-based decision-making during the ongoing data collection phase and ensures that desired quality standards are achieved.

Approximately 25% of the participants in every project location were given the role of "Monitor" based on their performance during the training. Each monitor was accountable for four to five schools. Each day, the monitor accompanied one data collector to the school and/or households to support the data collector and ensure that all processes were followed strictly. Another layer of quality control was provided by supervisors. Supervisors were mainly people such as ASER state team members and Pratham's local staff in respective states/districts, individuals belonging to external organizations, and those who had helped implement ASER surveys previously. These supervisors were responsible for the successful execution of assessment in the states, including all quality control processes (e.g., monitoring, conducting desk and field checks) (***Table E-1)***.

Table E-1. Quality control roles and responsibilities during data collection

Time Period	Data Collector	Monitor	Supervisor/ASER State Team
During data collection	Each data collector completes data collection in one school. This includes talking to the Head Master and explaining the evaluation, collecting school information, building rapport with Standard 2 students in a large group, tracking students for assessment, and completing the assessment with the tracked students. To accomplish the latter, they might also have to visit the households. Each data collector is assigned three to six schools, based on the size of the team.	Visits one school per day with one data collector and fills in the data collector booklet (marked with "Monitor" in type) for the visited school. 20–30% schools were monitored in this way.	Coordinates with monitors and data collectors by phone. Conducts monitoring visits to schools during data collection and fills out a monitoring form during field visits. This monitoring form captures whether data collectors follow correct protocol. Data collectors are spot trained by monitors when errors were found.
After data collection	No role.	Selected monitors assist with desk and field checks.	Performs desk and field checks. Once the hardcopies are submitted, they are thoroughly checked for completeness and correctness. Based on the feedback provided by the desk check, schools are selected for field checks. During field checks, the supervisors/ASER state team members speak to Head Masters and students and ask questions to assure that data collection was done following the correct protocol.

Table E-2 lists the number of schools surveyed, visited by monitors, quality controlled by supervisors, and desk and field checked.

Table E-2. Number of Schools Monitored and Checked

Program	Surveyed Schools		Schools Visited by Monitors		Schools Monitored by Supervisors		Schools Desk Checked		Schools Field Checked	
	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control
Scaling Up Early Learning Intervention—Uttarakhand	90	90	28	39	25	26	90	90	24	30
Scaling Up Early Learning Intervention—Chhattisgarh	60	60	18	18	12	16	60	60	10	10
Nurturing Early Literacy—Rajasthan	60	60	15	18	10	5	60	60	16	16
Nurturing Early Literacy—Karnataka	60	60	22	10	10	7	60	60	23	12
Nurturing Early Literacy—Maharashtra	70	70	21	16	16	13	70	70	20	14
Teacher Innovations in Practice—Uttar Pradesh	70	70	24	22	19	19	70	70	5	8
Start Early: Read in Time—Uttar Pradesh	70	70	27	24	23	30	70	70	12	5
Start Early: Read in Time—Odisha	60	60	19	20	15	14	60	60	22	23
RightToRead—Maharashtra	67	44	14	14	26	18	67	44	27	8
Total	**607**	**584**	**188**	**181**	**156**	**148**	**607**	**584**	**159**	**126**

Annex F: Additional Data on Student Characteristics

Table F-1 contains a summary of student demographic data collected at endline.

		Care Uttar Pradesh	Care Odisha	STiR Uttar Pradesh	EH Maharashtra	CMF Rajasthan	CMF Maharashtra	CMF Karnataka	R2R Uttarakhand	R2R Chhattisgarh
		Percent & SE	Percent & SE	Percent & SE	Percent & SE	Percent & SE	Percent & SE	Percent & SE	Percent & SE	Percent & SE
Female	Control	54.5% (3.5)	54.4% (2.5)	54.6% (2.1)	47.9% (6.1)	48.1% (2.4)	47.6% (1.9)	50.2% (1.7)	52.9% (2.4)	51.7% (2.2)
	Treatment	46.3% (2.5)	49.8% (3.3)	48.2% (2.4)	49.8% (2.1)	48.9% (0.4)	50.7% (2.3)	50.6% (1.1)	50.8% (1.7)	47.4% (2)
Breakfast before school	Control	85.1% (2.4)	92.2% (1.3)	92.7% (1.2)	86.5% (3.1)	87.5% (1.7)	88.8% (1.5)	89.2% (2)	91.5% (1.4)	93.9% (1.4)
	Treatment	89.5% (1.7)	96.3% (0.8)	91.9% (1.3)	84.2% (1.5)	89.1% (0.2)	87.1% (1.5)	93.3% (0.8)	95.6% (0.7)	93.6% (0.9)
Help with homework at home	Control	47.8% (2.8)	61.8% (3.5)	52.4% (2.5)	67.7% (3.4)	41.2% (2.8)	77% (2.2)	47.7% (3)	69.1% (2.6)	72.9% (2.1)
	Treatment	43.7% (2.9)	77.2% (2.6)	58.6% (3.1)	71.7% (1.7)	40.5% (0.4)	82.8% (2.2)	44.8% (2.4)	75.4% (2.1)	75.3% (2.1)
Extra reading material at home	Control	40.8% (3.7)	31.5% (3.4)	36.6% (2.8)	52.4% (4.8)	18.9% (2.2)	65.1% (2.6)	37.3% (2.9)	45.7% (3.3)	34.4% (2.7)
	Treatment	40.3% (3)	45.8% (3.5)	32.5% (3.1)	56.4% (2.6)	19.6% (0.3)	72.9% (2.6)	33.6% (2.2)	58% (2.2)	54.5% (3.4)
Absent from school in last week	Control	51% (3.3)	42.9% (4.3)	58.6% (2.4)	43.6% (3.4)	65.9% (2.5)	47.2% (2.3)	59.5% (2.1)	60.7% (3.5)	45.6% (2.8)
	Treatment	45.3% (2.4)	44.7% (3.7)	61.2% (2.6)	43.8% (2)	66.4% (0.4)	45.5% (2.7)	56.2% (2.6)	59.5% (2.4)	55.6% (3.1)
Someone read aloud at home	Control	36% (2.6)	46.4% (3.7)	33.7% (2.6)	59.3% (2.6)	28.5% (2.2)	72.7% (2.4)	44.7% (2.8)	57% (2.9)	51.4% (2.2)
	Treatment	37.8% (2.2)	58.8% (3)	37.8% (2.5)	57.6% (2.1)	28.3% (0.4)	76.9% (1.9)	36% (2)	65.7% (2.1)	61.5% (2.4)
Low SES	Control	34.7% (4.5)	36.7% (4.2)	35.2% (3.3)	38.8% (2.9)	55.4% (3.8)	1% (0.4)	44.3% (2.4)	31.4% (2.8)	42% (2.5)
	Treatment	38% (3)	19% (3.4)	36% (3)	29.8% (2.1)	63% (0.4)	1.3% (0.4)	46.2% (2.2)	24.6% (2.1)	37.2% (2.5)
Mid SES	Control	31.8% (3.5)	38.6% (3)	30.6% (2.6)	36.8% (2.3)	26% (2.4)	24.3% (2)	47.3% (1.9)	25.5% (2.4)	37.5% (2.3)
	Treatment	30.2% (1.9)	35.6% (3.1)	32.5% (2.1)	35.3% (2)	30.8% (0.4)	19.3% (1.7)	46.8% (2)	26.7% (1.9)	39.9% (2.1)
High SES	Control	33.5% (2.1)	24.7% (2.6)	34.2% (3.2)	24.4% (3)	18.6% (2.8)	74.7% (2)	8.4% (1.6)	43.1% (3)	20.5% (2.4)
	Treatment	31.8% (3)	45.4% (4.4)	31.6% (2.6)	34.8% (2.2)	6.3% (0.2)	79.4% (1.8)	6.9% (1)	48.7% (2.6)	22.8% (2.3)

Annex G: Comparing change in proportion of students at the Standard 2-level text in Treatment and Control Schools from Initial to Final Assessment

VARIABLES	Scaling up Early Reading Intervention - Uttarakhand	Scaling up Early Reading Intervention - Chhattisgarh	Nurturing Early Literacy - Rajasthan	Nurturing Early Literacy - Karnataka	Nurturing Early Literacy - Maharashtra	Teacher Innovations in Practice - Uttar Pradesh	Start Early: Read in Time - Uttar Pradesh	Start Early: Read in Time - Odisha	RightToRead - Maharashtra
Treatment	7.801**	10.23***	-1.826	2.820***	3.011	-2.236	1.455	5.140	-0.0867
	(3.052)	(3.039)	(2.161)	(0.945)	(5.696)	(1.972)	(2.114)	(3.409)	(0)
Constant	3.203*	11.10***	5.524***	0.501	10.29**	5.815***	3.402**	4.424***	1.034
	(1.920)	(1.830)	(1.523)	(0.541)	(4.853)	(1.178)	(1.325)	(1.661)	(0)
Observations	1,538	1,672	1,105	1,647	1,399	1,546	1,577	752	1,650
R-squared	0.012	0.016	0.001	0.005	0.001	0.001	0.000	0.004	0.000

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Change in Treatment significantly different from Control

All	Yes +	Yes +	No	Yes +	No	No	No	No	No

Treatment significantly different from Control after controlling for child and household characteristics

All	Yes +	Yes +		Yes +				No	

Annex H: Comparing Changes in the Proportion of Students at the Beginner Level in Treatment and Control Schools from Initial Assessment to Endline

Variables	Scaling Up Early Reading Intervention—Uttarakhand	Scaling Up Early Reading Intervention—Chhattisgarh	Nurturing Early Literacy—Rajasthan	Nurturing Early Literacy—Karnataka	Nurturing Early Literacy—Maharashtra	Teacher Innovations in Practice—Uttar Pradesh	Start Early: Read in Time—Uttar Pradesh	Start Early: Read in Time—Odisha	RightToRead—Maharashtra
Treatment	1.620	3.546	-6.100	-2.790	-0.0542	2.656	-6.165	-5.194	0.511
	(2.332)	(2.305)	(5.737)	(3.616)	(0.691)	(3.981)	(4.640)	(3.763)	(0)
Constant	-7.630***	-8.333***	-16.57***	-14.31***	-0.251	-18.01***	-16.26***	-4.661	-16.75
	(2.037)	(1.951)	(4.093)	(3.046)	(0.548)	(2.860)	(2.784)	(3.385)	(0)
Observations	1,538	1,672	1,105	1,647	1,399	1,546	1,577	752	1,650
R-squared	0.001	0.003	0.003	0.001	0.000	0.001	0.001	0.004	0.000

SEs in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Change in treatment significantly different from control

All	No	No	No	No	No	No	No	No	No

Annex I: Instruments (English Only)

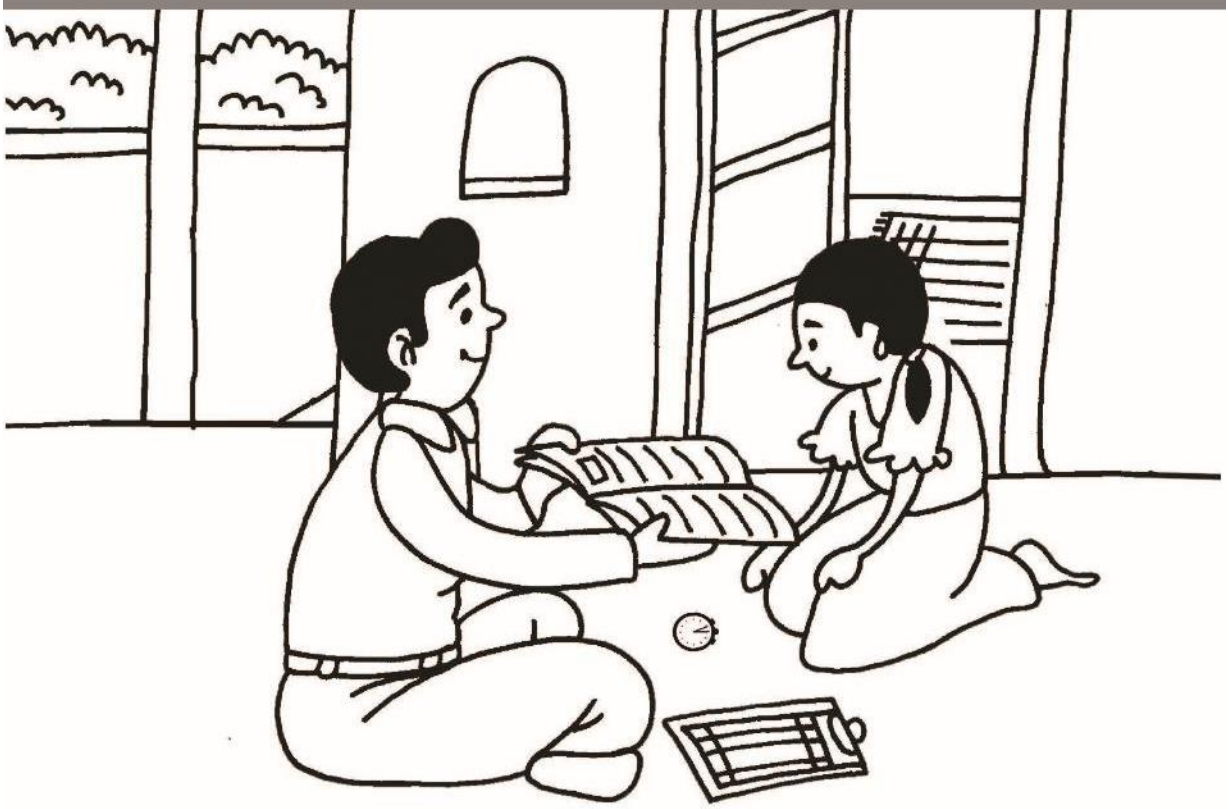
The English versions of the Student and Data Collector/Surveyor Assessment Booklets are included below. Instruments were developed for each project location for a total of six languages. Instrument versions for Gujarati, Marathi, Oriya, Kannada, and Hindi are available upon request.

Please contact:

Pratham Education Foundation
ASER Centre
B4/54, Safdarjung Enclave
New Delhi - 110029
Tel: +91 (011) 4602-3612

Early Grade Reading Assessment

Baseline Visit 2017



For child - give this booklet to the child.

ASER CENTRE, ASSESSMENT UNIT

USAID - BASELINE VISIT SEPTEMBER 2017

ASER Reading Test

Story

Salma is a little girl. She had a pretty doll. She loved playing with her doll. One day the doll fell from her hand to the floor. It broke into many pieces. Salma was very sad. She cried a lot. Her mother gave her another doll. Now she is happy again.

Para

Ravi is a boy.
He has many friends.
He loves to draw.
He does not like to sing.

Para

My village is very big.
It has many houses.
It also has a shop.
The bus stops in my village.

ASER Reading Test

Letter

b	s	o
k	m	
y	r	h
t	x	

Ask the child to recognize any 5 letters. At least 4 must be correct.

Word

ring	bad
ball	
cold	king
clap	foot
fan	
girl	crow

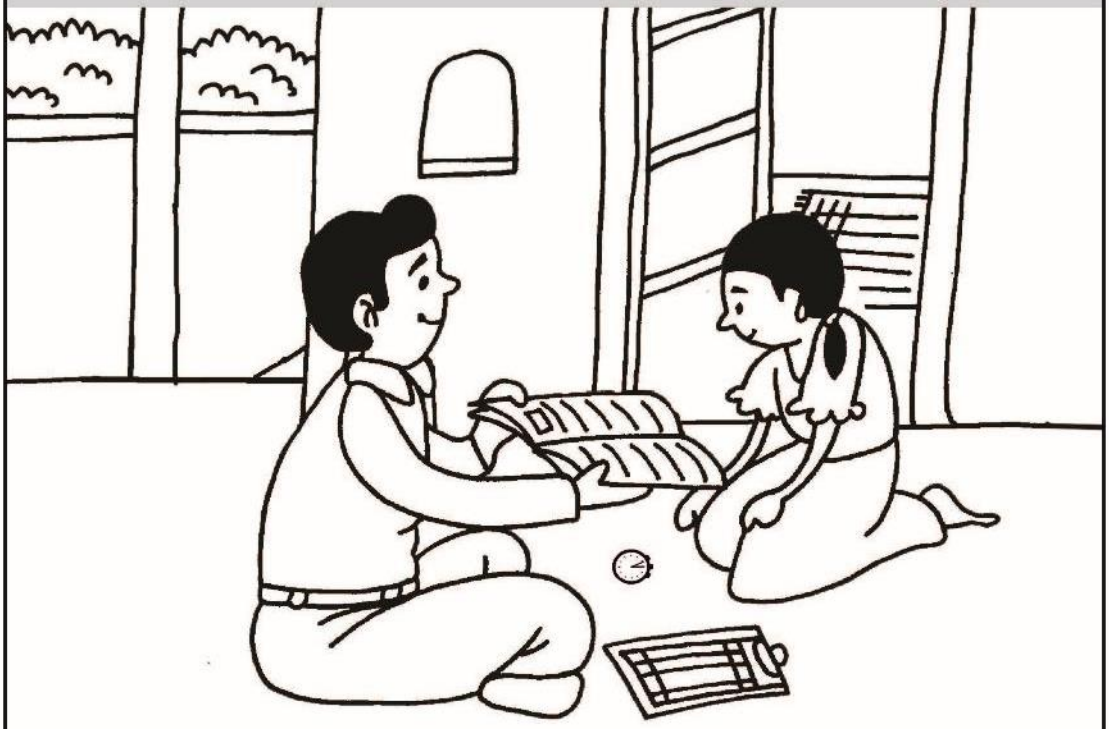
Ask the child to read any 5 words. At least 4 must be correct.

ORF Reading and Comprehension

Story

Gita goes to school every day. Her mother gave her a book. The book had a red cover. Gita read the book every morning on her way to school. Gita learned many new words from the book. That made her teacher happy. The teacher gave Gita another book. It had nice stories. She showed it to all her friends.

Early Grade Reading Assessment
Baseline Visit 2017



For surveyor – don't give this booklet to the child.

Survey Information

Type *(Pick One)*	Surveyor	Monitor

State name		SCHOOL CODE	
School name			
Surveyor ID		Date of survey	
Surveyor name		Mobile number	

School Information

Visit only the schools given by the ASER Team. Meet the headmaster of the school. In the absence of the headmaster, meet the senior-most teacher of the school.

Respondent's name		**Mobile number**	
Designation of Respondent *(Pick One)*	**Headmaster**		**Teacher**

Enrollment in class 2 *(Fill from the enrollment register)*	**Boys**	**Girls**
Attendance in class 2 *(Count on your own)*	**Boys**	**Girls**

Assent from children

(Before testing the child, take assent following the process given below.)

Note to surveyor: After visiting the school and explaining the purpose of your visit to the Headmaster go to the classroom where grade 2 children are sitting. Play one of the games given in the manual or any other game with the children.

Before starting the assessment establish a playful and relaxed rapport with each child through a short conversation. The child should perceive the assessment almost as a game to be enjoyed rather than a test.

To take assent, read aloud to the child slowly and clearly ONLY the text given in the box.

- Good morning. My name is _____ and I live in _____.
- We are going to play some reading games. I am going to ask you to read out aloud some letters, words and stories. I will also ask you some simple questions about the story.
- Using this watch, I will see how long it takes you to read. (Show the stopwatch to the child)
- **This is NOT a test and it will not affect your grade at school.**
- I will also ask you some questions about your house, for instance, some of the things that are there at your house.
- You do not have to participate if you do not wish to.
- Do you want to participate? Are you ready to get started?

Scoring Sheet

(Use one sheet per child.)

Part 1 – Child Information

CHILD NUMBER				**Child's full name**	
Age		**Gender**	**M** F	**Father's full name**	
Village name or ward name				**Name of hamlet in the village or the address/ landmark/any other indicator of the house**	

Did the child assent to participate in the assessment: Yes ☐ **No** ☐

(If verbal consent is not obtained, thank the child and move on to the next child)

Part 2 – Child Questionnaire

Q1. Did you eat any food/breakfast before you arrived at school today?
a) No ☐ b) Yes ☐
c) Don't know ☐ d) No response ☐

Q2. Generally, does someone at home help you with your homework/studies?
a) No ☐ b) Yes ☐
c) Don't know ☐ d) No response ☐

Q3. Apart from your school books, are there any other books, newspapers, etc. for you to read at home?
a) No ☐ b) Yes ☐
c) Don't know ☐ d) No response ☐

Q4. Were you absent from school any day in the last week?
a) No ☐ b) Yes ☐
c) Don't know ☐ d) No response ☐

Q5. Generally, does anybody at home read aloud stories, poems, etc. to you?
a) No ☐ b) Yes ☐
c) Don't know ☐ d) No response ☐

Q6. Do you have an electricity connection at home?
a) No ☐ b) Yes ☐
c) Don't know ☐ d) No response ☐

Q7. Do you have a toilet at home?
a) No ☐ b) Yes ☐
c) Don't know ☐ d) No response ☐

Q8. Generally, where does drinking water come from at your home?
a) Tap in house ☐ b) Community tap ☐
c) Well ☐ d) Handpump ☐
e) Water tank ☐ f) River/lake/stream ☐
g) Others _____ ☐ h) Don't know ☐
i) No response ☐

Q9. Generally, how is food cooked at your home?
a) Electric stove (heater, induction) ☐
b) Gas/biogas stove ☐
c) Kerosene stove ☐
d) Firewood stove ☐
e) Coal stove ☐
f) Others _____ ☐
g) Don't know ☐
h) No response ☐

Q10. Do you have the following items in your home?

	No	Yes	Don't know	No response
a) Mobile phone				
b) Television				
c) Computer				
d) Radio				
e) Refrigerator				
f) Car				
g) Motorcycle				
h) Cycle				

Q11. Generally, what language do you speak with your family at home?

Part 3 – ASER Reading Level

(Mark at the highest level)

Beginner	Letter	Word	Paragraph	Story

Part 4 : ORF Reading and Comprehension

Reading

Surveyor tells the child: "Read this story aloud but carefully. When you finish, I will ask you some questions about what you have read.
When I say "Begin," read the story as best as you can. If you come to a word you do not know, go on to the next word.
Put your finger on the first word of the story. Ready? Begin."

- **Start the timer as soon as you say "Begin".**
- **When the timer reaches 0, tell the child to "Stop".**
- **If a child hesitates while reading or stops on a word for 3 SECONDS, say "Go on."**
- **If the child does not read a single word or does not read a single word correctly on the first sentence then finish the exercise here and put a tick mark (✓) in the first box given below the story.**

(/) **Mark any incorrect or unread words with a slash (/).**
(Ø) **If you accidentally mark a word wrong or if the child self-corrects an incorrectly spoken word, then circle (Ø) it. Count such words as correctly read words.**
(]) **Mark the final word read with a bracket (]) at the end of 60 seconds.**

Story for reading *(Child reads out of his/her booklet)*	Number of correct words
Gita goes to school every day*. Her mother gave her a book. 1 2 3 4 5 6 7 8 9 10 11 12	☐/12
The book had a red cover. Gita read the book every morning on her 13 14 15 16 17 18 19 20 21 22 23 24 25 26 way to school*. 27 28 29	☐/17
Gita learned many new words from the book*. 30 31 32 33 34 35 36 37	☐/8
That made her teacher happy. The teacher gave Gita another book*. 38 39 40 41 42 43 44 45 46 47 48	☐/11
It had nice stories. She showed it to all her friends*. 49 50 51 52 53 54 55 56 57 58 59	☐/11

Comprehension

Surveyor tells the child: "Now I am going to ask you a few questions about the story you just read. Try to answer the questions as well as you can. You can provide your answers in whichever language you prefer."

After the child is finished reading, REMOVE the story from in front of the child.

- **If the child does not read a single word or does not read a single word correctly on the first sentence, do not ask any comprehension questions.**
- **Ask the child only the questions related to the part of the story read. For every part, ask the question only if the child has read all the text up till the (*).**
- **If the child does not provide a response to a question after 10 seconds, mark the child at 'no response' and continue to the next question.**
- **Do not ask the question more than twice.**
- **If a child says "I don't know," mark the child at 'incorrect'.**

Question [Answer]	N.A.	Correct	Incorrect	No response
Where does Gita go every day? [School]				
What did Gita do on her way to school? [Read] or [Read the book]				
What did Gita learn from the book? [Words] or [New words]				
Who gave another book to Gita? [Teacher]				
Why did Gita show the book to all her friends? [Gita's teacher gave it to her.] or [The book had nice stories.]				

If the child does not read a single word or does not read a single word correctly on the first sentence then finish the exercise here and put a tick mark (✓) in the box:

If the child reads all the words in less than a minute then note the *remaining* number of seconds (out of 60):

Last word read (write the number):

Total number of correctly read words:

Annex J: Scope of Work

Introduction and Background

USAID/India seeks to improve the quality and accuracy of measurement approaches of its projects so that all USAID/India early grade reading activities are able to measure and report progress towards Goal 1 of the USAID Education Strategy consistently and reliably.

March 2017, the E3/Ed Evidence Team submitted a desk review in response to a technical assistance request from USAID/India. The purpose of the work was to conduct a review of 13 early grade reading projects' reading assessment methodologies, tools, data and analyses. The key finding was that only one of 13 reviewed activities had assessment methodology and tools that were compatible with the Global Count methodology[12] and an additional two projects that could be adjusted to be compatible. The remaining 10 assessment approaches were not compatible or the quality of the assessment results could not be ascertained. A recommendation from the review is that the Mission implement a modified EGRA or ASER using a standardized approach that aligns with accepted standards across all projects. An initial assessment will be conducted in September/October 2017 with a follow-up assessment conducted at the end of the school year in March/April 2018. For West Bengal schools of Right To Read, the study will be done in January and July 2018.

The primary purpose of this research is to be able to include the reading assessment results of USAID India's reading projects in the Global Count, as well as to be able to compare results across projects. USAID would like RTI to partner with Pratham/ASER Centre to determine what modifications are needed, if any, to the ASER assessment and reporting methodology so that the assessment can serve as a common measure of reading fluency across projects in USAID's portfolio and so that the data can be used to report into the Global Count. The data will also be used to report Indicator E.S.1-1 for these six projects "Percent of learners who demonstrate reading fluency and comprehension of grade level text at the end of grade 2 with USG assistance."

The six activities captured in Table 1 will be included in this research study.[13]

Table 1: Snapshot of activities

ACTIVITY	Period of Performance	Dates of School Year
School Excellence Program (SEP)	July 2014 –July 2018	June to May
Start Early: Read in Time	July 2014 –July 2018	April to March
Teacher Innovations in Practice (TIPs)	October 2014 – September 2018	April to March
Right to Read	September 2015 – September 2017	Maharashtra - June to April West Bengal – January to July
Nurturing Early Literacy	October 2015 – September 2019	Various; in 3 states (see ***Annex C***)
Scaling up Early Reading Interventions (SERI)	September 2015 – September 2020	April to March

Study methodology

- **Study methodology**: Measure increases in student performance over the course of one school year in treatment and control schools (beginning and end of the 2017

[12] Measuring and Communicating Progress towards Goal One. http://pdf.usaid.gov/pdf_docs/PA00KR71.pdf

[13] School Excellence Program (SEP) was later removed from the study.

school year).

- **Student sample**: A random sample of boys and girls, from grade 2. Sample size is 10 boys and 10 girls per school.
- **School sample**: See Table 2 and reference Annex D: Note on Scope and Sampling (60 treatment and 60 control schools per intervention). See ***Table 2*** and reference ***Annex K***.
- **Geographic locations:** Eleven locations covered across 6 projects – Uttarakhand, Chhattisgarh, Maharashtra, Gujarat, Uttar Pradesh, Odisha, Karnataka, Rajasthan, and West Bengal.
- **Instrumentation**: For the global count, we propose to assess Grade 2 children using an adapted ASER instrument. The adapted ASER instrument will include an oral reading fluency subtest for all children. The oral reading fluency subtest will include a second fluency passage. This passage will be different to the standard ASER grade 2 long text reading passage and will also have 5 comprehension questions. At least one of the five questions will be an inferential question. The additional passage used to measure fluency will be an alternate passage from ASER's passage pool and will be of comparable difficulty. A short student questionnaire will accompany the assessment. Standard ASER testing protocols will be used for the ASER portion of the assessment (adaptive assessment). For the fluency passage, all children will be given 60 seconds to read the fluency passage and answer corresponding comprehension questions as related to how far they read into the text.
- **Assessment methodology**: The assessment will be conducted one-on-one, by trained assessors, in school, in an appropriate setting with minimal distractions. Standard ASER protocol is that two assessors administer the assessment. One assessor administers the assessment and the second assessor marks the assessment. Data will be collected on paper.
- **Language of assessment**: The language of instruction of the intervention will be used for the assessment. It will be important to note that language of instruction varies across projects. In Right to Read for example, the intervention is focused on teaching children to read English as a second language. It is important to note that results across languages cannot be compared. Languages covered – 1) Hindi, 2) Marathi, 3) Gujarati, 4) Oriya, 5) Kannada, and 6) English as second language for Right to Read Project (used in Maharashtra and West Bengal).

Detailed Work Plan

Implementation

Research design. The first step in the research will be to document the research design and evaluation methodology in an Evaluation Design and Implementation Plan (EDIP). This is a short report that documents the research plan as well as key assumptions. This report will be led by RTI in partnership with Pratham. The EDIP supports fidelity of design and is a key piece of documentation.

Adapt ASER instrument and protocols. The ASER instrument will be adapted to include a second reading passage that will be used to measure oral reading fluency. Five comprehension questions will be added to the assessment. The second reading passage will be taken from ASER's passage pool and will be of equivalent difficulty to the standard ASER passage. For this assessment, ASER will develop 5 comprehension questions for each language. In addition, instructions will be added to the standard ASER administration protocols to reflect the addition of these new subtasks.

Pilot instrument revisions. The instrument will be piloted with approximately 200 children. The goal is to achieve at least 150 non-zero scores on the reading fluency passage. The focus of the pilot will be to test the administration of the assessment (total duration, switching the order of the assessments, team configurations, inter-rater reliability implementation), and the 5 comprehension questions. Additional piloting requirements will be determined in collaboration with Pratham. Pratham will conduct the psychometric analysis to review the instrument's properties. The tool will be finalized in consultation with RTI and shared with USAID in advance of the larger data collection. The findings from the psychometric analysis and instruments revisions will be documented in the EDIP.

Training. RTI will participate in the training provided by Pratham on the ASER assessment. Pratham will lead the training of the data collection teams. During the training, assessors will be tested to assess the reliability and accuracy of their test administration (Assessor Accuracy Measure). The AAM will be conducted at the end of the training and used to identify any assessors that do not meet administration standards. Assessors scoring less than 80% will not be asked to participate in the data collection. The training will accommodate at least one trip to a nearby school where assessors will be able to administer the instrument with children and receive coaching and feedback. Following the pilot, feedback is provided to refine test protocols. Pratham will update administration protocols and training manuals. To build capacity of project staff, staff from each project will be asked to attend the training. At the end of the training, project staff will be capacitated to serve as observers of future data collection efforts and monitors of data quality. Prior to the end-of school data collection, a refresher training will be held for assessors.

Data Collection. Data collection, analysis and reporting will be led by Pratham. Initial data collection will be conducted in September with the second data point collected at the end of the school year in 2018. The end of the school year varies for each project, see ***Annex C***.

Data Analysis. Pratham will lead the data analysis and reporting. An initial results report will be drafted for USAID following the data collection in September. This report will follow standard ASER reporting but will also include information that can be used for the Global Count for each project and reporting on Indicator E.S. 1-1 for each project. Following the second, end-of the year data collection, a final results report will be submitted.

Capacity Building. Project staff will be invited to attend the ASER/EGRA training. Staff will have the opportunity to be trained on proper test administration and quality control. Travel and lodging will be at USAID's/ the project's expense. Following the data analysis, RTI will lead a benchmarking workshop in New Delhi for each language (with relevant projects) and the Ministry. Travel and lodging costs for project staff are not included in this budget.

Table 2: Minimum Sample Size: All large projects

PROJECT	NUMBER REACHED	GEOGRAPHY	# OF SAMPLED SCHOOLS	# OF SAMPLED STUDENTS
School Excellence Program (SEP)	150,000 children (Grades 1-8)	Gujurat	60T + 60C (only Gujrati medium schools)	2400
Start Early: Read in Time	100,000 children (Grades 1-4)	Uttar Pradesh and Odisha	60T + 60C in Uttar Pradesh 60T + 60C in Odisha	4800
Teacher Innovations in Practice (TIPs)	564,000 children (Grades 1-8)	Uttar Pradesh	60T + 60C	2400
Right to Read	1,000,000 children, (Grades 1-8)	Maharashtra (Nagpur, Latur, Solapur, Pune, Osmanabad and Jalgaon districts) and West Bengal	60T + 60C Maharashtra 60T + 60C in West Bengal	4800
Nurturing Early Literacy	93,334 children (primary grades)	Rajasthan, Karnataka, and Maharashtra	60T + 60C in Rajasthan (only schools where Bodh's program is being implemented) 60T + 60C in Karnataka 60T + 60C in Maharashtra	7200
SERI	330,000 (Grades 1-5)	Uttarakhand and Chhattisgarh	60T + 60C in Uttarakhand 60T + 60C in Chhattisgarh	4800
TOTAL	**2,237,334**		**1320 (660T + 660C)**	**26,400**

*60 treatment schools and 60 control schools

Report Writing

Evaluation Design Implementation Plan. RTI will lead the drafting of the EDIP, with key inputs from Pratham. The EDIP will document the research methodology and key assumptions. This will be drafted within 10 days of project start or before any data collection takes place. The EDIP will be finalized following the pilot. The EDIP will serve to document findings from the pilot as well as the results of the psychometric analysis.

Final instrument. Following the pilot, the final adapted ASER instrument, single-form, will be shared with USAID prior to data collection. Relevant updated protocols and manuals will accompany the final instrument.

Initial report. The results of the initial data collection will be submitted in a summary report. The report outline will follow standard ASER reporting but will include data that can be used for the Global Count and to report on Indicator E.S. 1-1 for each project. A presentation of results will also be scheduled with USAID in New Delhi. RTI will attend remotely. As West Bengal is on a different school schedule (January – December) the base line report will be finalized following the completion of the West Bengal data collection in January, 2018.

Final report. Following the final data collection at the end of the 2017 school year a final results report will be submitted. Following submission of the final report, an in-person presentation will be scheduled with USAID in New Delhi. RTI will attend remotely. As West Bengal is on a different school schedule (January – December) the base line report will be finalized following the completion of the West Bengal data collection in July, 2018.

Deliverables

- Evaluation Design Implementation Plan
- Final instruments
- Initial report
- In-country presentation (Pratham; RTI remotely)
- Final report
- In-country presentation (Pratham; RTI remotely)

Annex K: Revisions to SOW

This brief note lists down some of the critical decisions about sampling and would need agreement from USAID and RTI before we proceed with sampling. This includes inline comments from ASER, USAID and RTI. These revisions were agreed to by all parties on 11 August 2017.

Research Design:
The proposed evaluation approach is called a difference-in-difference. Typically, it is used in a cross-sectional evaluation design where the intervention gain minus the control gain gives a year-on-year learning impact. It is proposed that the same approach is used to measure impact between September 2017 and March 2018. A new sample of control and treatment schools will be selected for the baseline.

Limitations:
There are some limitations with this design:

- The students assessed in grade 2 intervention schools have already received the intervention in grade 1. As a result, balancing the baseline control and treatment groups using covariates such as socio-economic status (SES) becomes impossible as the students in the intervention of equivalent SES to control students should be demonstrating better reading skills.
- The analysis approach only measures impact over a 6 month period.
- Measuring the USAID 1.1 indicator of the improved percent of students reading fluently cannot be calculated in the usual way because this is usually done by a cross-sectional analysis of just the treatment group over 1-year. Making the same calculation over a 6 month period entangles the learning gain from with the gain typically seen from 6-months of learning.
- As with the 1.1 indicator measuring the global count of number of improved students cannot be calculated by its usual method of (like the 1.1 indictor) assessing the improvement seen by the treatment group over a one-year cross-sectional period.

Due to these limitations, where possible it is recommended that control and treatment schools in the project original baseline sample be included in the 2017 baseline sample. These schools were selected pre-intervention and consequently, these schools can be balanced using the original baseline student scores (no matter what instrument was used) and SES data collected in the 2017 baseline. This creates a balanced baseline which can then be used for an **end line only** analysis; a t-test which compares the treatment and control reading scores. This analysis will be equivalent to a *balanced* difference-in-difference analysis with the additional benefit of being able to measure student reading gains from grade 1 and grade 2 combined. Because the sample size for this method is smaller it might lack statistical power for impact evaluation, but because the global count and 1.1 indicators do not use precision, so this is a non-issue. It is recommended that both the global count and the 1.1 indicator use an end line treatment versus control difference; as the analysis is balanced this method will produce the same results as a baseline-end line treatment gain.

Response from ASER
Agreed that the gains will be measured only over a 6 month period. However, having said that, the measurement is being done over the most productive time during the school year. Learning loss is often observed, when children come back from summer vacations (July). It takes time for children to settle down and we have observed that the learning gain observed during a full school year is close to that observed during September-March of the school year.

Ideally, for each program, there should have been an end line at the end of grade 1 (baseline for grade 2) and then another end line at the end of grade 2. This approach would enable disentangling the learning gain in grade 1 from that in grade 2. Since, this was not done, RTI suggests including the historical treatment and control schools in the current sample. Our approach, from the start, has been that this exercise should be clean with what is being measured clearly defined. If we had a clean sample of treatment and control schools, RTI's suggestion would make perfect sense, and in fact reduce our work considerably. However, from what has been shared with us it seems that each program assessment was done independently and there is no evidence of a uniform research design across these interventions.

In particular, note that:

1. Programs are at different points in their intervention cycle and have even been discontinued in some districts and/or added in new districts. For example, for the CARE program in UP, the historical baseline was done in 87 treatment schools in 3 districts. And, there is no historical control group. Today, the program is in 5 districts. Our sample size is 60 treatment and 60 control schools per state per program. If all 87 schools are included, two districts will not be covered in the sample. If additional schools are included from the new districts, that will have cost and time implications.
2. A control group was not selected for each intervention. A case in point is the CARE program in UP. If we stick with the original treatment schools, and select a new control now, the argument of a balanced baseline doesn't really hold.
3. In some cases even a baseline measurement was not done. For e.g., English Helper program in Maharashtra and West Bengal. Therefore, for this program the only estimate would be from the current exercise.
4. Even where a baseline was done and/or a control selected, we are not tracking the same set of children. Therefore, comparing the end line with the original baseline is essentially an aggregated cross-sectional analysis, the limitations of which are well known.

One could look at each program/state separately and devise a sampling strategy on a case-by-case basis with the objective of retaining the historical treatment and control schools in the sample. However, even though the sampling is being done so as to evaluate at the program level in each state, the sample will be pooled to estimate the effectiveness of the USAID interventions as a whole, and as far as possible we want a uniform research design for that.

Finally, the point of precision not being an issue is a little confusing – if this exercise is going to provide estimates of USAID effectiveness as well as proportion of grade 2 children at grade competency, precision will be an important concern. The requirements for the estimate may be different than for impact evaluation – and these can be worked out – nevertheless issues of precision cannot be overlooked.

We strongly feel that this exercise should not be contaminated by historical decisions. The reason for undertaking this exercise is precisely that there was no consistent evaluation approach followed across the interventions. On the other hand, we understand the need to use the historical sample to be able to say something about the learning gains over the entire duration of the interventions. Therefore, we have suggested including the historical sample as an additional sample in the end line.

RTI: Thank you for consideration of our sampling approach.

2. Project level decisions

a. We propose that Right to Read project implemented by English Helper Technologies be excluded from this evaluation for the following reasons:
    - The project is very thinly spread across many districts of Maharashtra and West Bengal. If we sample 60 treatment schools per project per state, then for each district (within a state) we will get very few schools in our sample.
    - The project aims to improve children's reading and comprehension in English (as second language) whereas other projects work for the first language.
    - The schools covered under this program are extension schools unlike the other projects where projects have been running since last couple of years.

USAID/RTI position: Include Right to Read on the basis that USAID requires data on Right to Read for ES Indicator 1.1. We understand that the data from Right to Read will not be directly comparable to the other projects and that is not the intention, but rather being able to contribute to the global count as well as indicator ES 1.1.
ASER: May be then we need to select a few districts where most schools are covered. Also, for West Bengal we need to check the school calendar – some schools follow Jan-December

b. For School Excellence Program we propose to cover only Gujrati medium schools. If we intend to include other languages the sample size for treatment schools will increase – we would need 60 treatment schools per language.

USAID/RTI position: Only include Gujrati medium schools for the School Excellence Program. Gujarat was later removed from the SOW.

c. For STiR Education's program (Teacher Innovations in Practice) we will cover all 10 districts in Uttar Pradesh. 2 districts (of these 10) – Varanasi and Rae Bareli are also covered in our independent evaluation of STiR's program.

RTI position: Understood.

d. For Nurturing Early Literacy program in Sirohi district, Rajasthan 2 separate organisations are working. Are these 2 different interventions? If yes, we would need 60 treatment schools per intervention. Or we can include the intervention that includes most schools.

RTI position: To be confirmed by USAID.
USAID: Yes, there are two different interventions operating in the Sirohi district, one by Bodh Shiksha Samiti and the other by Room to Read. USAID suggests including Bodh's intervention for this assessment under the Nurturing Early Literacy Project, since Room to Read is already included in this study under the SERI project.

3. School level sampling decisions

a. For any intervention, irrespective of the number of schools covered within a state – we will sample 60 treatment and 60 control schools. If the intervention runs in more than 1 district in a state, the treatment and control schools will be sampled proportionately based on the number of intervention schools in each district with an exception of excluding the districts which have very few intervention schools (for logistical purposes).
b. To sample the treatment and control schools Probability Proportional to Size (PPS) sampling technique on grade 2 enrollment numbers will be used. The sample of treatment schools will be drawn out from school list shared by USAID India and should only include schools where interventions will definitely be running in year 2017-18. We would need USAID India to share school lists in a more standard format so that sampling can be done using the list.

c. Control schools will be sampled from the same blocks as the treatment schools. A criteria of matching enrollment numbers and a few other characteristics from DISE (District Information System for Education) will be used to sample control schools.
d. RTI addition: Where there are baseline schools from prior data, they will be included in the sample for this planned assessment. Including the original schools, for which we have the pre-intervention scores, will keep open the possibility of measuring the project's impact. In addition, the inclusion of these original schools will also allow us to balance the sample based on relative school performance prior to intervention.
    - ASER: We need to check whether these baseline schools are still part of the intervention. We have asked USAID India to share the school list for 2017-18
    - RTI: Understood.
    - ASER (8/11/17): We strongly feel that this exercise should not be contaminated by historical decisions. The reason for undertaking this exercise is precisely that there was no consistent evaluation approach followed across the interventions. On the other hand, we understand the need to use the historical sample to be able to say something about the learning gains over the entire duration of the interventions. Therefore, we have suggested including the historical sample as an additional sample in the end line.
    - RTI: Understood. Thank you for consideration of our sampling approach.
e. RTI addition: School and learner demographic information will also be collected through a student questionnaire to allow for propensity matching.

Note: Even if the sample designs are different (i.e. PPS was not previously used) you can have data with schools from different designs. For example, SRS and PPS – their weight calculations will be different but can still be analyzed together.

4. Child level sampling decisions

a. In each school we will randomly sample about 20 children (10 boys and 10 girls) in grade 2 for the baseline assessment. In case a randomly sampled child is not present in the school on the day of visit, then the child will be replaced by another student from the same class who is present. In the end line visit (in March/April 2018) we will track all sampled children from the baseline visit in the community/villages if they are not present in school on the day of visit.

b. According to the MDES calculations using ASER 2016 data for the state of Uttar Pradesh – the sample size proposed in the point above will get us somewhere between 0.2 and 0.25 MDE.
Following assumptions are used for MDE calculation:
Data from ASER 2016
State = Uttar Pradesh (UP)
Grade 2 children enrolled in government schools
Intraclass correlation at the village correlation to proxy for school level – this is not such a bad assumption since most children in grade 2 within a village are likely to be in the same school

The calculations were done separately for girls and boys.
Numbers used:

Proportion Boys in grade 2, in Govt schools in UP, at story level in ASER test = 0.0384; SD=0.192; rho=0.078
Proportion girls in grade 2, in Govt schools in UP, at story level in ASER test = 0.0252; SD=0.157; rho=0.1

Here is what we get:

MDE=0.2 * SD

Boys – Minimum number of schools = 118 with 13 boys per school
Girls – Minimum number of schools = 118 with 18 girls per school

MDE=0.25 * SD

Boys – Minimum number of schools = 120 with 6 boys per school
Girls – Minimum number of schools = 115 with 7 girls per school