# All Children Reading–Asia

# Analysis of Early Grade Reading Assessment in India

# Initial Data Collection and Assessment Report

# Table of Contents

# List of Figures

# List of Tables

# List of Acronyms and Abbreviations

| | |
|---|---|
| AAM | assessor accuracy measure |
| ASER | Annual Status of Education Report |
| CmF | Centre for microFinance |
| cwpm | correct words per minute |
| DISE | District Information System for Education |
| EGR | early grade reading |
| EGRA | early grade reading assessment |
| ES | Education Strategy |
| ORF | oral reading fluency |
| PCA | principal components analysis |
| PTR | pupil-teacher ratio |
| PPS | probability proportional to size |
| RTE | Right of Children to Free and Compulsory Education Act |
| SD | standard deviation |
| SDG | Sustainable Development Goal |
| SE | standard error |
| SES | socio-economic status |
| SRS | simple random sampling |
| SSA | Sarva Shiksha Abhiyan |
| STIR | Schools and Teachers Innovating for Results |
| TBD | to be determined |
| USAID | United States Agency for International Development |
| USG | United States Government |

# 1      Executive Summary

The US Agency for International Development (USAID)/India seeks to improve the quality and accuracy of measurement approaches of its projects so that all USAID/India early grade reading activities can measure and report progress toward Goal 1 of the USAID Education Strategy consistently and reliably. The purpose of this evaluation activity is to provide data that will allow the Mission to report toward the Global Count as well as towards the reading indicator *the percent of learners 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)*. This report presents the results and findings from the initial data collection and assessment of five large education projects within USAID/India's portfolio.

## 1.1      Research Background

In March 2017, the USAID Bureau for Economic Growth, Education and Environment/Education Department 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 1 of 13 reviewed activities had assessment methodology and tools that were compatible with the Global Count methodology and an additional 2 projects had data 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 that followed from the review was for the Mission to implement a standardized assessment approach to align accepted standards for research design and instrumentation across all projects.

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. These projects are captured in *Table 1* below. Section 3 of the report provides detailed information on the research design and the instruments used to assess students' reading ability.

**Table 1.      Snapshot of USAID/India-funded early grade reading projects included in evaluation**

| Project Name | Implementing Partner | Period of Performance |
|---|---|---|
| **Scaling Up Early Reading Intervention** | Room to Read | September 2015 – September 2020 |
| **Nurturing Early Literacy** | Centre for microFinance | 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 |
| **Right to Read** | English Helper | September 2015 – September 2017; extension in Maharashtra 2017–2019 |

## 1.2    Research Design

The evaluation design to measure gains in reading performance will be a difference analysis whereby individual student gains will be calculated (endline score minus initial assessment score). Each project location's impact will be 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 two assessments. Initial data were collected at the beginning of the school year and will be collected again at the end of the school year by retesting the same students across treatment and control schools. *It is important to note that although the data collection may be referred to in evaluation planning documentation as a "baseline" it is not a true baseline. Many projects are in either 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.*

The evaluation sampled and collected data from Standard 2 students. Data collection started September 15, 2017, and concluded by the end of September for the majority of project locations. Data collection was delayed for two projects in Maharashtra (Nurturing Early Literacy and Right to Read) due to delays in gaining data collection permissions. Data collection in these locations was completed by October 14, 2017.

The ASER/modified EGRA instrument was administered to a total of 14,370 Standard 2 students randomly selected from within 1,191 government primary schools (607 treatment and 584 control) from 7 states and 31 districts in India. Participating schools were randomly selected to create a sample that would be representative of selected project districts. Sample design took into consideration the geographical spread and language of the interventions. In all, nine project locations were assessed independently. 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

The ASER reading assessment indicates the percentage of children who are reading at beginner level, letter level, word level, and text level.[1] During administration of the ASER reading assessment, each child was marked at the highest level at which they could read comfortably. This means that a child marked at Standard 2-level text can read at all lower levels of the tool. For letters and words, the child was expected to read at least four out of five given letters/words correctly. For Standard 1- and Standard 2-level text, the child was expected to read the text fluently (as sentences rather than as a string of words) with three or fewer mistakes.

As noted above, multiple languages were assessed depending on the projects' language of instruction. The Right to Read 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.

*Table 2* shows the percentage of students reading at Standard 1-level and Standard 2-level text in treatment and control schools for each project location. Similar to the results of the modified EGRA the results of the ASER reading assessment were wide ranging. The lowest percentage of students reading at Standard 2-level in the treatment group was 0.1% with a high of 53.9%. The range of Standard 2-level text readers in the control group ranged from

---

[1] See Table 9.Overview of ASER assessment reading tasks for definition of levels.

0.0% to 57.1%. The results of tests of equivalence by project location, for the initial assessment, are also provided for the Standard 2-level text indicator under the column effect size. These results show that the magnitude (as standard deviation) of difference between control and treatment for three projects (Scaling Up Uttarakhand, Scaling Up Chhattisgarh, and Start Early Odisha) is large, but as will be explained later in the executive summary, it is difficult to draw any definite conclusion. Note that where the number of students in the sample reading Standard 2-level text was too small, the effect size was suppressed.

The results in *Table 2* are sorted by the percentage of students reading at text level. The order of project locations mirrors the results of the modified EGRA presented in the following section in *Table 3.*

**Table 2.      Percent of children reading at text level across treatment groups**

| Project-Location | Sample Size | | Standard 1-Level Text (% of students) | | Standard 2-Level Text (% of students) | | | Text Level (% of students reading at text level–combined) | |
|---|---|---|---|---|---|---|---|---|---|
| | Treatment (T) | Control (C) | T | C | T | C | Effect size | T | C |
| Nurturing Early Learning – Maharashtra | 814 | 656 | 25.7% | 27.2% | 53.9% (2.7) | 57.1% (3.1) | 0.06 | 79.6% | 84.3% |
| Scaling Up Early Reading Intervention – Uttarakhand | 974 | 707 | 23.9% | 10.4% | 20.3% (2.2) | 11.1% (2.0) | 0.22 | 44.2% | 21.5% |
| Scaling Up Early Reading Intervention – Chhattisgarh | 932 | 950 | 23.7% | 12.9% | 17.0% (2.5) | 6.8% (1.4) | 0.25 | 40.7% | 19.7% |
| Start Early: Read in Time – Odisha | 497 | 408 | 12.9% | 4.2% | 22.3% (2.5) | 14.7% (2.3) | 0.22 | 35.2% | 18.9% |
| Teacher Innovations in Practice – Uttar Pradesh | 896 | 869 | 2.7% | 2.9% | 6.1% (1.1) | 4.9% (1.4) | 0.07 | 8.8% | 7.8% |
| Nurturing Early Learning – Karnataka | 1,039 | 783 | 5.4% | 4.0% | 1.6% (0.4) | 1.4% (0.4) | (-) | 7.0% | 5.4% |
| Start Early: Read in Time – Uttar Pradesh | 946 | 826 | 2.6% | 3.0% | 3.8% (0.7) | 3.6% (1.2) | 0.05 | 6.4% | 6.6% |
| Nurturing Early Learning – Rajasthan | 666 | 591 | 3.3% | 2.7% | 0.3% (0.0) | 1.2% (0.5) | (-) | 3.6% | 3.9% |
| Right to Read – Maharashtra | 1,084 | 752 | 0.8% | 0.6% | 0.1% (0.0) | 0.0% (0.0) | (-) | 0.9% | 0.6% |

*Standard errors in parentheses
(-) Suppressed due to small sample size

*Figure 1* summarizes the percentage of children at each ASER reading level by project location. The vast majority of children across treatment and control are reading at beginner and letter levels. The majority of children in the control and treatment schools in Nurturing Early Literacy – Maharashtra schools are text level readers. There are three project locations where a significant number of children in treatment schools are text level readers: Scaling Up Early Learning Intervention in Uttarakhand and Chhattisgarh and Start Early in Time – Odisha. After controlling for child and household characteristics for both project locations

treatment scores are significantly higher from control of the Scaling Up Early Reading Intervention project.

**Figure 1.    ASER results summary**
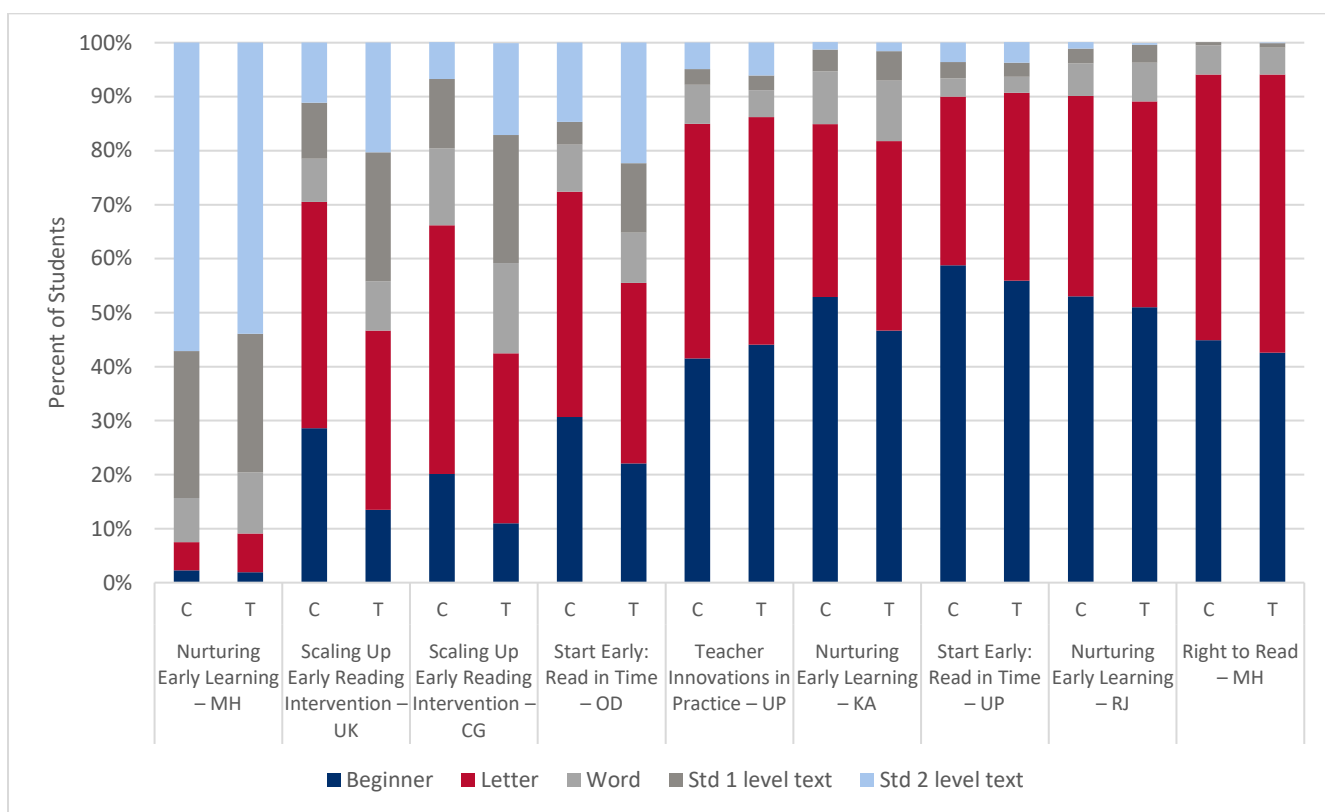


## 1.4    Oral Reading Fluency and Reading Comprehension Findings

The outputs from the modified EGRA are shown in **Table 3**. Oral reading fluency (ORF) measured in correct words per minute (cwpm) and reading comprehension measured as percent of questions answered correct. The EGRA results show a wide range of mean ORF and reading comprehension scores across projects but also across project locations. ORF means ranged from 0.7 cwpm to 35.3 cwpm in the treatment group, and 0.6 cwpm and 30.2 cwpm in the control group. Similarly, a wide range of scores was evident for reading comprehension. In treatment, the range was between 0.2% correct to 44% correct, and 0.1% correct in the control to 40.5% correct in the control. With the exception of Nurturing Early Learning – Maharashtra, ORF and reading comprehension scores were low across project locations and treatment groups.

**Table 3** also features the percent of students who could not read a single word correct, or scored zero, on the ORF subtask. The percent range of zero scores was between 1.6% and 93.3% in the treatment group and 5.2% to 95% in the control. Projects with relatively high ORF means had a lower percentage of students scoring zero in the ORF subtask. The inverse is also true: a lower ORF mean is associated with a greater percentage of students scoring zero. Although the percent of children scoring zero is not an indicator for this activity, it is a beneficial metric to review in conjunction with ORF means to understand and contextualize an improvement in reading gains.

*As mentioned above, this is not a "traditional" baseline and so careful interpretation of the reading scores in context is necessary.*

The last three columns of **Table 3**, average socio-economic status (SES), the language of assessment, and home language help add context to the data. The last column provides the

percentage of children in the sample that speak the same language at home as the language of assessment. In five of nine project locations, the majority of children reported speaking at home a different language than the language of assessment. The SES categories were created across all the project locations, such that the SES categories by project location in *Table 3* can be compared. Generally, the higher the average SES the higher the control group reading averages.

*What could explain the difference between scores across projects?*

- Variation in language complexity and orthography
- Variation in district-level SES and other contextual district differences

*What could explain the difference between treatment and control scores within projects?*

- Difference in SES

- Pre-treatment differences; Treatment effect; potential impact of Standard 1 intervention

**Table 3.    Modified EGRA summary results by project name and location**

| Project-Location | Oral Reading Fluency Mean | | Reading Comprehension % Correct | | Percent Zero Scores | | Average SES | Language of Assessment | Majority Home Language (% of sample) |
|---|---|---|---|---|---|---|---|---|---|
| | T | C | T | C | T | C | | | |
| Nurturing Early Literacy – Maharashtra | 35.3 (33.1, 37.6) | 30.2 (28.5, 31.9) | 44% (39.9%, 48.0%) | 40.5% (38.0%, 43.0%) | 1.6% (0.8%, 3.4%) | 5.2% (3.5%, 7.7%) | HIGH | Marathi | Marathi (97.1%) |
| Scaling Up Early Reading Intervention – Uttarakhand | 15.7 (13.3, 18.1) | 7.3 (5.4, 9.3) | 13.7% (10.9%, 16.5%) | 7% (4.6%, 9.4%) | 31.1% (25.9%, 36.9%) | 61.2% (54.3%, 67.6% | MIDDLE | Hindi | Hindi 76.2%) |
| Scaling Up Early Reading Intervention – Chhattisgarh | 14.6 (12.8, 16.5) | 6.6 (4.7, 8.5) | 9.8% (8.1%, 11.4%) | 3.9% (2.3%, 5.4%) | 27.1% (22.5%, 32.3%) | 52% (43.1%, 60.8%) | MIDDLE | Hindi | Chhattisgarhi (83.8%) |
| Start Early: Read in Time – Odisha | 10.8 (8.1, 13.5) | 5.8 (4.3, 7.3) | 10.3% (7.1%, 13.4%) | 6.5% (4.6%, 8.3%) | 48.7% (41.5%, 56%) | 62.9% (54.8%, 70.4%) | MIDDLE | Oriya | Other (55.8%) |
| Teacher Innovations in Practice – Uttar Pradesh | 3.7 (2.5, 4.9) | 3.1 (1.9, 4.4) | 3.3% (2.1%, 4.5%) | 3.4% (1.9%, 4.9%) | 81.8% (75.9%, 86.5%) | 82.4% (76.9%, 86.8%) | LOW | Hindi | Hindi (88.3%) |
| Nurturing Early Literacy – Karnataka | 2.6 (2.0, 3.1) | 2.1 (1.3, 2.9) | 2.1% (1.2%, 2.9%) | 1.0% (0.4%, 1.6%) | 65.9% (61.3%, 70.3%) | 71.9% (65.4%, 77.7%) | LOW | Kannada | Kannada (59%) |

| Project-Location | Oral Reading Fluency Mean | | Reading Comprehension % Correct | | Percent Zero Scores | | Average SES | Language of Assess-ment | Majority Home Language (% of sample) |
|---|---|---|---|---|---|---|---|---|---|
| | T | C | T | C | T | C | | | |
| Start Early: Read in Time – Uttar Pradesh | 2.4 (1.6, 3.1) | 2.4 (1.2, 3.5) | 2.8% (1.8%, 3.8%) | 2.5% (0.8%, 4.2%) | 89.3% (85.5%, 92.3%) | 89.2% (85.3%, 92.2%) | LOW | Hindi | Hindi (70.7%) |
| Nurturing Early Literacy – Rajasthan | 1.7 (1.6, 1.7) | 1.3 (0.8, 1.9) | 0.6% (0.6%, 0.6%) | 0.7% (0.4%, 1.1%) | 80.9% (80.4%, 81.4%) | 86.8% (81.6%, 90.6%) | LOW | Hindi | Grasiya (40.8%) |
| Right to Read – Maharashtra | 0.7 (0.4, 0.9) | 0.6 (0.2, 1.0) | 0.2% (0.1%, 0.3%) | 0.1% (-0.1%, 0.3%) | 93.3% (91%, 95%) | 95% (90.2%, 97.5%) | HIGH | English | Marathi (65.8%) |

95% confidence intervals in parentheses

## 1.5 ORF Mean per ASER Reading Category

*Table 4* presents mean ORF scores per each of the ASER reading categories by project location. The sample for each project location includes both treatment and control schools. The table indicates that higher ASER reading levels are associated with higher ORF means, as expected. In most project locations, students were reading at the beginner or letter level, reading 0 cwpm. At the Standard 2 level as well as at the Word and Standard 1 levels, there is a large amount of variation in the cwpm across the different project locations. The Standard 2 reading level ranges from 23 cwpm in the Nurturing Early Literacy program in Rajasthan to 46 cwpm in the Right to Read project in Maharashtra. Even within projects, the students who were able to read connected text (categorized at the Standard 1 or Standard 2 levels) spread across a vast range of reading abilities in terms of fluency.

**Table 4.    ORF mean scores per ASER reading category by project name and location**

| Project-Location | N | Oral Reading Fluency Mean | | | | |
|---|---|---|---|---|---|---|
| | | **Beginner** | **Letter** | **Word** | **Standard 1** | **Standard 2** |
| Nurturing Early Literacy – Maharashtra | 1470 | 0 (0) | 4.82 (5.38) | 12.81 (7.9) | 25.6 (11.06) | 43.05 (16.32) |
| Nurturing Early Literacy – Rajasthan | 1257 | 0 (0.03) | 0.64 (1.76) | 8.17 (6.86) | 17.07 (8.34) | 23.18 (8.82) |
| Nurturing Early Literacy – Karnataka | 1822 | 0.13 (1.09) | 1.13 (2.07) | 6.66 (5) | 15.39 (9.11) | 38.24 (27.38) |
| Start Early: Read in Time - Odisha | 905 | 0.03 (0.3) | 0.7 (1.96) | 6.95 (4.93) | 15.68 (11.55) | 33.56 (14.98) |
| Start Early: Read in Time – Uttar Pradesh | 1772 | 0.03 (0.37) | 0.46 (2.53) | 5.85 (6.36) | 17.96 (9.99) | 41.57 (21.13) |
| Scaling Up Early Reading Intervention – Chhattisgarh | 1882 | 0.04 (0.38) | 1.57 (3.05) | 9.69 (6.08) | 22.29 (9.91) | 36.82 (16.04) |
| Scaling Up Early Reading Intervention – Uttarakhand | 1681 | 0.08 (0.87) | 1.68 (3.72) | 8.77 (7.49) | 22.53 (12.11) | 38.3 (19.84) |
| Right to Read – Maharashtra | 1836 | 0 (0) | 0.21 (1.43) | 9.63 (7.31) | 15.01 (8.47) | 46 (0) |
| Teacher Innovations in Practice – Uttar Pradesh | 1765 | 0.03 (0.44) | 0.75 (2.31) | 5.85 (7.3) | 17.75 (7.09) | 38.15 (19.8) |

Standard errors in parentheses

## 1.6 Important Consideration for Interpretation of Results

The baseline should always be assessed for equivalence (or balance) between the control and treatment groups to make sure that the two groups are comparable prior to the intervention. In other words, they are in similar geographic regions, with similar resource levels, same language of instruction, and most importantly, have students with similar average reading scores and similar demographic characteristics. In this way, we can be confident that the difference between control and treatment scores at endline can be attributed to the intervention and the two groups were homogeneous—an "apples to apples" comparison. *Table 5* presents differences in scores between control and treatment groups

for ORF and reading comprehension subtasks (see **Table 2** for effect sizes between control and treatment groups for the ASER Standard 2 reading level).

An effect size greater than 0.25 indicates that a difference between control and treatment groups does not satisfy baseline equivalence.[2] However, as explained for **Table 3**, this difference cannot necessarily be fully attributed to demographic differences between the control and treatment groups, as the difference can also be attributed to gains from the intervention in Standard 1. So these balance tests are inconclusive by research design. These differences were also statistically adjusted using a covariate regression model, controlling for average differences between the groups such as wealth and help with reading at home. These results, which are shown in **Annex A,** made little impact on the baseline equivalence other than for the Start Early: Odisha program, where the effect size went down from 0.34 to 0.18 for reading fluency.

**Table 5.  ORF and reading comprehension balance across control and treatment groups by project location**

| Project-Location | Treatment | Oral Reading Fluency | | Reading Comprehension % Correct | |
|---|---|---|---|---|---|
| | | Average | Difference (effect size in parentheses) | Average | Difference (effect size in parentheses) |
| Nurturing Early Literacy – Maharashtra | Treatment | 35.3 | 5.1 (0.28) | 44.0% | 3.5% (0.12) |
| | Control | 30.2 | | 40.5% | |
| Scaling Up Early Reading Intervention – Uttarakhand | Treatment | 15.7 | 8.4 (0.52) | 13.7% | 6.7% (0.34) |
| | Control | 7.3 | | 7.0% | |
| Scaling Up Early Reading Intervention – Chhattisgarh | Treatment | 14.6 | 8.0 (0.57) | 9.8% | 5.9% (0.38) |
| | Control | 6.6 | | 3.9% | |
| Start Early: Read in Time – Odisha | Treatment | 10.8 | 5.0 (0.34) | 10.3% | 3.8% (0.20) |
| | Control | 5.8 | | 6.5% | |
| Teacher Innovations in Practice – Uttar Pradesh | Treatment | 3.7 | 0.6 (0.06) | 3.3% | -0.1% (0.01) |
| | Control | 3.1 | | 3.4% | |
| Nurturing Early Literacy – Karnataka | Treatment | 2.6 | 0.5 (0.07) | 2.1% | 1.1% (0.13) |
| | Control | 2.1 | | 1.0% | |
| Start Early: Read in Time – Uttar Pradesh | Treatment | 2.4 | 0.0 (0.00) | 2.8% | 0.3% (0.02) |
| | Control | 2.4 | | 2.5% | |
| Nurturing Early Literacy – Rajasthan | Treatment | 1.7 | 0.4 (0.07) | 0.6% | -0.1% (0.02) |
| | Control | 1.3 | | 0.7% | |
| Right to Read – Maharashtra | Treatment | 0.7 | 0.1 (0.03) | 0.2% | 0.1% (0.03) |

---

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

Detailed ASER reading assessment and modified EGRA results are presented for each project location in Section 6. The report begins with background information on the Government of India's initiatives and enacted legislation since 2000 and USAID/India's current education portfolio to provide local context to our research objective. Section 3 provides information on research design, sampling and importantly, limitations of the research. A brief discussion of the development and piloting of instruments is explored in Section 4. Information on training methodology and data collection is provided in Section 5. Following Section 6, the Results Section, a short conclusion in provided in Section 7. The last section, Next Steps, outlines what will follow this initial report: the benchmarking workshop to set reading benchmarks for each language.

# 2    Introduction

## 2.1    Background and Local Context

Sarva Shiksha Abhiyan (SSA) was operationalized in 2000–2001 as a Government of India flagship program to achieve universal elementary education in the country. This program seeks to build new schools in areas that do not have adequate school facilities. It also strengthens existing school infrastructure through provision of additional classrooms, toilets, and drinking water, etc. Schools with inadequate numbers of teachers are provided with additional teachers. Furthermore, capacity of existing teachers is built through training and academic support at cluster, block, and district levels (SSA Shagun, 2017).

In addition to the SSA, the Constitution (Eighty-Sixth Amendment) Act, 2002 inserted Article 21-A in the Constitution of India to provide free and compulsory education to all children aged 6–14 years as a fundamental right. The Right to Education (RTE) Act, 2009, the legislation envisaged under Article 21-A, came into effect April 1, 2010. Among other things, the RTE Act mandates free and compulsory education through completion of elementary education in a neighborhood school. The Act also makes provisions for a child who is not enrolled in school to be admitted to an-age appropriate class. Norms and standards relating to the pupil-teacher ratio (PTR), buildings, and infrastructure are also laid down in the RTE Act (Ministry of Human Resource Development, 2015).

As a result, in India, more than 95% of children between the ages of 6 and14 years old are now enrolled in school—and more than 97% of 7–10-year-olds enroll in primary school since 2007.[3] School infrastructure has also improved in last few years. According to the Annual Status of Education Report (ASER) Centre survey of 2016, 53% of schools are now compliant with the RTE Act's PTR norms, compared to 38.9% in 2010, and 74.1% and 68.7% of schools have drinking water and useable toilet facilities available, resepectively (ASER Center, 2016).

Despite the success in increasing enrollment and improving infrastructural facilities, learning outcomes have not kept pace with increased school access. ***Table 6*** shows the proportion of children in Standard 5 who can read a Standard 2-level text on the ASER reading assessment by the type of institution in which children are enrolled. According to ASER Centre (2016(, only 47.8% of students in Standard 5 can read a Standard 2-level text. Students who do not acquire these basic reading skills in earlier grades usually lag behind in higher grades. Even in Standard 8, 27% of students are struggling to read a Standard 2-level text (ASER Centre, 2016). As shown in the table, tracking the reading trends over time shows that India's education system's ability to impart basic reading skills seems to be weakening over time.

---

[3] Based on the data published in successive ASER survey reports. These reports are available on the ASER Centre website –
http://www.asercentre.org/Keywords/p/236.html

**Table 6.    Reading levels over time: All India (rural) students in Standard 5 who can read a Standard 2-level text**

| Year | Enrolled in Government Schools | Enrolled in Private Schools | Enrolled in Government and Private Schools* |
|------|-------------------------------|----------------------------|--------------------------------------------|
| 2010 | 50.7 | 64.2 | 53.7 |
| 2012 | 41.7 | 61.2 | 46.9 |
| 2014 | 42.2 | 62.5 | 48.1 |
| 2016 | 41.6 | 62.9 | 47.8 |

*This is the weighted average for students enrolled in government and private schools.
Source: ASER, 2016, p. 52

On January 1, 2016, the United Nations 17 Sustainable Development Goals (SDGs) to transform the world by 2030—adopted by world leaders in September 2015 —officially came into force. SDG 4 sets out the goal to ensure "inclusive and equitable quality education for all and promote lifelong learning." These goals can only be achieved by improving the foundational skills of reading and numeracy for all students.

## 2.2    USAID/India's Efforts to Improve Reading

One of the three major goals laid out in USAID's Global Education Strategy (USAID/India, 2017) is to contribute to the improvement of the reading skills for 100 million children in the primary grades. By focusing on innovation and strategic partnerships USAID is building on India's dynamic entrepreneurial climate where public and private sectors are already rising to meet the country's development challenges. To meet this global target, USAID/India is working with the government and local partners to support innovative early grade reading interventions. The five USAID/India supported EGR interventions that are included in this evaluation represent the Mission's commitment to identify, test, and scale-up innovations that sustainably advance educational outcomes in India. The interventions are varied and multifaceted and span delivery of teaching and learning materials, support of teacher networks, and assisted technology. USAID/India also has a key focus on building capacity to improve pedagogy and teaching by working with teachers, headmasters and other education officials. ***Annex B*** provides details of each intervention.

This research activity aims to provide a standardized measure of reading outcomes that can be counted toward the goal of 100 million children as well as provide information on project effectiveness. Data on how well reading is improving helps is the foundation for further improvement.

# 3    Research and Sample Design

## 3.1    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, 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 towards 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.

## 3.2 Research Design

Of the 13 USAID/India-supported early grade reading (EGR) projects, 5 projects are included in this evaluation report (see *Table 7* below). All 13 projects could not be evaluated due to funding limitations; therefore, projects were selected based on size and scope. The selected projects represent those that have the largest scopes and, as such, have the biggest potential impact on USAID's Goal One target. The evaluation design aims 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.

The sample of control and treatment schools for the initial data collection was drawn in August 2017. This study is longitudinal at the student level, meaning the same students are 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 have been (or will have been by the final assessment) exposed to a USAID education project (treatment) and Standard 2 students who have not been exposed to any USAID-funded education project (control). The longitudinal design of this study warrants a difference analysis whereby student reading skill gains will be calculated at the two time points. This will allow us to generalize the results for the Standard 2 population per project location for both the initial and final data collections as well as evaluate each project location's impact based on the gains achieved in treatment schools as compared with control schools.

As summarized in *Table 7* below, seven states are included in the evaluation: Uttarakhand, Chhattisgarh, Rajasthan, Uttar Pradesh, Odisha, Karnataka, and Maharashtra. In total, nine project locations were evaluated. Sampling was designed to get representative estimates for reading performance of Standard 2 students in each of these project locations. See Section 3.3 below 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. The language of instruction varies across projects (see *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 evaluated project's 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 | Room to Read | Uttarakhand | Almora, Champawat, Dehradun, Udham Singh Nagar (4) | Hindi |
| | | Chhattisgarh | Baloda Bazar, Raipur (2) | |
| Nurturing Early Literacy | Centre for microFinance | 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, Varansi (10) | Hindi |
| Start Early: Read in Time | CARE | Uttar Pradesh | Bahraich, Balrampur, Gonda, Hardoi, Shravasti (5) | Hindi |
| | | Odisha | Mayurbhanj (1) | Oriya |
| Right to Read | English Helper | Maharashtra | Nagpur, Latur, Solapur, Pune, Osmanabad and Jalgaon (6) | English |

## 3.3    Sampling Design

As previously mentioned, the five USAID/India-funded EGR projects included in this evaluation are spread across 7 states and span 31 districts in India. Due to language and other cultural, socio-economic, 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 9 project locations. The sampling design provides representative estimates for reading performance of Standard 2 students in each of these project locations. The final sample by project and state is presented in *Table 8.* Detailed district-level sample information is available in *Annex C*.

**Table 8.     School and student sample by project location**

| Program | Number of Schools * | | | Student Sample in Treatment Schools | | | Student Sample in Control Schools | | |
|---|---|---|---|---|---|---|---|---|---|
| | Treatment | Control | Total | Boys | Girls | All** | Boys | Girls | All |
| Scaling Up Early Reading Intervention – Uttarakhand | 90 | 90 | 180 | 473 | 497 | 974 | 333 | 368 | 707 |
| Scaling Up Early Reading Intervention – Chhattisgarh | 60 | 60 | 120 | 472 | 455 | 932 | 458 | 489 | 950 |
| Nurturing Early Literacy – Rajasthan | 60 | 60 | 120 | 331 | 335 | 666 | 299 | 292 | 591 |
| Nurturing Early Literacy – Karnataka | 60 | 60 | 120 | 505 | 534 | 1,039 | 394 | 389 | 783 |
| Nurturing Early Literacy – Maharashtra | 70 | 70 | 140 | 410 | 404 | 814 | 343 | 313 | 656 |
| Teacher Innovations in Practice – Uttar Pradesh | 70 | 70 | 140 | 438 | 449 | 896 | 420 | 440 | 869 |
| Start Early: Read in Time – Uttar Pradesh | 70 | 70 | 140 | 498 | 433 | 946 | 428 | 394 | 826 |
| Start Early: Read in Time - Odisha | 60 | 60 | 120 | 240 | 247 | 497 | 189 | 207 | 408 |
| Right to Read – Maharashtra | 68 | 44 | 112 | 548 | 536 | 1,084 | 370 | 382 | 752 |
| **Total** | **607** | **584** | **1191** | **3914** | **3871** | **7828** | **3234** | **3274** | **6542** |

\* See Section 3.3.1 below about sampling of additional schools.
\*\* The sum of boys and girls does not always add up to "All" since gender was not recorded for some students.

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. Since schools vary in size, schools were sampled using probability proportional to size (PPS)[4] 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 sampled per project location. Treatment schools were sampled from the lists provided by 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 socio-economic status (SES) in treatment schools. However, there are no such secondary data 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 is provided below.

1. Sample treatment schools.

---

[4] PPS is a sampling technique in which the probability of selecting a sampling unit (school in our case) is proportional to 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.

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. 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, due to 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. Nearly 50% of schools reported enrollment of fewer than 20 students, and on the day of data collection 76% of schools had fewer than 20 students in attendance (see ***Annex D***).

### 3.3.1  Sampling of Additional Schools

Originally, 120 schools (60 treatment and 60 control) were to be sampled in each project location. However, during data collection, a real-time online system 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 due to low attendance and registration in schools. Therefore, it was decided to sample additional schools in some project locations.

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. Teacher Innovations in Practice – Uttar Pradesh: An additional 10 treatment and 10 control schools were sampled

3. Start Early: Read in Time – Uttar Pradesh: An additional 10 treatment and 10 control schools were sampled

4. Start Early: Read in Time – Odisha: An additional 10 treatment and 10 control schools were sampled

5. Nurturing Early Literacy – Maharashtra: An additional 10 treatment and 10 control schools were sampled

6. Right to Read – Maharashtra: An additional 10 treatment and 10 control schools were sampled

### 3.3.2  Incomplete Data Collection

Project locations where all sampled schools were not surveyed are:

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. Right to Read – 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.

3. Right to Read – West Bengal: State-level training for this project location was successfully completed in January 2018; however, data collection could not begin due to challenges in obtaining permission from the local government.

4. School Excellence Program – Gujarat: State-level training for this project location was successfully completed from September 11–13, 2017, in Surat, Gujarat; however, data collection could not begin due to challenges in obtaining permission from the local government.

## 3.4    Research Limitations

As mentioned above, differences in individual students' scores will be used to help measure each project location's impact based on the gains achieved in treatment schools as compared with control schools. However, there are important interpretations and limitations to consider when reviewing project outcomes when data are not collected prior to the start of project implementation.

The main purpose of a true baseline is to assess "balance" between treatment and control groups pre-intervention; in other words, the treatment and control groups should be comparable and, therefore, the final assessment gains of the treatment group over the control group can be attributed to the intervention. If at initial assessment the mean scores are not statistically similar and individual student learning gains to some extent are attributable to student scores at baseline, then the difference between the control and treatment gains cannot be fully attributable to the intervention. As mentioned in the Executive Summary, it was not possible to conduct the baseline balance assessment.

It is important to note that this assessment was conducted with Standard 2 students; almost all students in sampled treatment schools were part of intervention programs beginning in Standard 1. Consequently, the treatment mean scores described in Section 6 Results are a combination of achievement and students receiving 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. If school matching achieved balance, then these initial results show impact of the interventions in Standard 1 on reading fluency and comprehension at the start of Standard 2.

Additional limitations:

- Due to a variety of delays, the initial data collection did not begin until mid-September 2017. Since the final assessment will be implemented before the end of the school year and exams in February/March 2018, the evaluation will measure impact over a 4 to 5-month period. ASER's previous research has shown that September–March is the most productive time of the school year with the largest learning gains being observed in this period. The delay was further compounded in a few project locations because some school districts did not allow field teams to collect data from schools in their jurisdiction. Data collection was completed in mid-October for these schools.

- Ideally, control schools should be matched to the learning levels in treatment schools. However, there are no such secondary data available at the student level or even at the school level. The only information available at the school level is about infrastructure from the District Information System for Education (DISE), which makes balancing between treatment and control schools on the basis of student outcomes impossible. Since the initial assessment was done after one year of intervention (Standard 1 year), we are not truly able to assess the balance between control and treatment 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 risk entangling the learning gain attributed to the intervention with any gains typically seen from school learning.

- The same students will be retested using the same instrument; therefore, they will have familiarity with the structure and content of the assessment when being tested at the final assessment. However, if a child's reading abilities improve by final assessment, the child may be able to read further in the passage (which will be new content not read at initial assessment) and receive additional questions.

- Spillover effects may be seen. Control schools are sampled from the same blocks (district sub-divisions) as treatment schools to ensure closer matching. It is possible that teachers and cluster and block-level officers meet and interact 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 as well as provided with the exact days that data collectors would visit the schools to assess students. These notifications might result in the Hawthorne effect, meaning 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. Through investigation by ASER it was determined that this school was unsuitable for the sample and was removed from the analysis.

# 4    Assessment Instruments

Standard 2 students in sampled schools were assessed with the ASER reading assessment and a modified version of the EGRA. Standard ASER reading assessment protocols were followed during the ASER portion of the assessment.[5] For the ORF and comprehension subtasks, standard EGRA protocols were also followed.[6]

The assessment also included a short student questionnaire that collected information about socio-economic status, attendance, and other student demographics.

Informed assent was collected from students prior to beginning the assessment. All data were collected on paper; data collectors used stop watches as necessary for timed subtasks. English versions of the Instrument booklets are included in ***Annex E***.

## 4.1    Global Count – ASER

The ASER Survey is an annual household-based, citizen-led assessment of learning outcomes in rural India. By design ASER is a "floor" test that aims to evaluate students' early reading ability. The reading assessments were developed taking into account the state-mandated curriculum for each state. The content of the reading assessment (i.e., the selection of words, the length of sentences and reading passages) was aligned to the Standard 1- and 2-level textbooks in each state. At the letter level, recognition of only simple letters is assessed (secondary forms of letters and conjoint letters are not usually part of Standard 1 curriculum in most states). At the word level, simple one- and two-syllable words, commonly used every day and appropriate for Standard 1 are included. In the development

---

[5] For more information on ASER administration protocols, see
http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/aserassessmenttasks.pdf
[6] For more information on EGRA administration protocols, see http://shared.rti.org/content/early-grade-reading-assessment-egra-toolkit-second-edition

of Standard 1- and 2-level passages, orthography-specific indicators such as the use of simple letters, secondary representations of letters, and conjoint letters have been considered along with sentence and passage length. Vocabulary used in the reading passages is aligned to the state-mandated curriculum for appropriateness. **Figure 2** is a sample of the ASER reading test in Hindi.

**Figure 2.    ASER reading test in Hindi**



The ASER reading assessment categorizes learners 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. **Table 9** provides an overview of the ASER assessment reading tasks.

**Table 9.    Overview of ASER assessment reading tasks**

| Subtask | Scoring Methodology | Assesses… |
|---|---|---|
| **Standard 2-level text** | Student can read at story (Standard 2) level if the student reads the text fluently with 3 or fewer mistakes | Indicates the accuracy and fluency with which students read aloud a Standard 2-level passage comprising 7–10 sentences and about 60 words |
| **Standard 1-level text** | Student can read at paragraph (Standard 1) level if the student reads the text fluently with 3 or fewer mistakes | Indicates the accuracy and fluency with which students read aloud a Standard 1-level passage comprising 4 sentences and about 20 words |
| **Word reading** | Student is asked to read any 5 words (of 10); at least 4 must be correct to be marked at word level | Indicates the accuracy with which students read aloud out of a list of 10 one- or two-syllable words |
| **Letter identification** | Student is asked to read any 5 letters (of 10); at least 4 must be correct to be marked at letter level | Indicates the accuracy with which students read aloud out of a list of 10 letters |

## 4.2    Indicator E.S.1-1 – Modified EGRA

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 ORF subtask. For this evaluation, the EGRA portion only included ORF and reading comprehension subtasks (see **Figure 3**

below for the Hindi example of this subtask). The ORF subtask was an additional reading passage that was timed using a stopwatch.

**Figure 3.    EGRA ORF and comprehension subtasks in Hindi**



The ORF passages for each language were drawn from ASER's story bank of Standard 2 passages. Each child was allocated 60 seconds to read the passage. Data collectors marked words that the child read incorrectly or could not read and recorded the total number of words attempted and read correctly at the end of the reading task. The ORF passage was divided into 5 parts and there were 5 comprehension questions based on each part of the passage. Based on where in the passage the child reached in the 60 seconds allotted, the appropriate number of comprehension questions were asked to the child. For students who could not read even one word correctly in the first sentence, the reading as well as the comprehension tasks were discontinued. ***Table 10*** summarizes the EGRA subtasks used in this study.

**Table 10.    Overview of EGRA direct assessment reading tasks**

| Subtask | Scoring Methodology | Assesses… |
|---|---|---|
| **Oral reading fluency** | Students are asked to read aloud a grade-level short story printed on a page. (*Timed subtask: 60 seconds*) | Automatic and word recognition |
| **Reading comprehension** | Students are asked to verbally respond to five oral questions (four literal and one inferential) based on the short story and on how far the child read into the text. (*Untimed subtask*) | Comprehension |

## 4.3    Student Questionnaire

A short student questionnaire was also administered at the end of the assessment. The questions gathered important information about SES and other student demographics, such as student absenteeism.

The following questions, such as whether the student has electricity and a toilet at home or access to certain consumer goods, if the student had breakfast before school, and if there are additional reading materials at home, were used as proxies to create a SES index using principal components analysis (PCA). PCA describes the association across variables, and subsequently, variables that are highly associated were used to create a SES index for each project. The SES index was then split into three equal groups to classify students as having low, mid, or high SES. Because each PCA model was run separately for each project, the SES distribution presented for each project represents a relative measure of wealth within the project population.

# 5    Pilot, Training, and Data Collection

It is important to ensure that the instruments used for the evaluation were functioning properly in order to collect reliable and valid data. A pilot was conducted between August 29 and September 2, 2017. The pilot was completed by experienced ASER data collectors in 7 states and 8 districts. For the pilot, approximately 2,400 students in Standard 2 were assessed. Additional details on the pilot and pilot findings are presented in ***Annex F***. Psychometric analysis was also performed and summary results are also included in ***Annex F***. Pilot findings informed administration and teaming protocols; instruments were not revised.

Following the pilot, data collectors were trained using a two-tier, cascade training model. A master training was held centrally September 7 to 9, 2017, in Jaipur, Rajasthan. During this training, 26 master trainers were trained on the administration of the assessment. The following week these master trainers were deployed to state training venues, where they trained the surveyors (data collectors).

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 289 participants were selected as data collectors to participate in the field work. An additional 80 participants served as data quality monitors.

Each field team was made up of the following people: data collector, monitor, and supervisor/ASER state team. Detail on the training of data collectors and data collection dates are provided in ***Annex G***. Information on quality control measures is included in ***Annex H***.

# 6    Initial Assessment Results

This section presents the main findings from the initial assessment. Student demographics are presented, followed by students' performance on the ASER reading assessment and the ORF and comprehension subtasks by each project location.

The initial results show the proportion of students at each level for the ASER reading assessment, and the mean scores for ORF and reading comprehension by control and treatment. The effect size Cohen's *d* has been provided, which is calculated by 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 intervention impact. Cohen attributed a small, medium, and large effect size to values of 0.2, 0.5 and 0.8, respectively (Cohen, 1992).

As previously discussed in Limitations, Section 3.4, this assessment is not a true baseline (e.g., it was not conducted prior to the start of intervention), which has major implications for the findings presented below. The main limitation is that it is 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, SES data collected provide some evidence that this was not achieved across all project locations.

## 6.1 Scaling Up Early Reading Intervention Project

Room to Read 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, learners were assessed in Hindi in four districts in Uttarakhand and two districts in Chhattisgarh.

### 6.1.1 *Scaling Up Early Reading Intervention Project – Uttarakhand*

Of the students assessed in Uttarakhand for the Scaling Up Early Reading Intervention project, 53% in the control group were girls and 47% were boys. In the treatment group, 51% of students were girls and 49% were boys. The majority of students assessed indicated they speak Hindi (about 77%), which is the language the assessment was conducted in, while a small proportion responded that they speak a language other than one of the six language options listed in the assessment.

*Figure 4* shows that students in control schools were comparable with those in treatment schools in terms of their SES. The largest proportion of students from both groups were categorized as high SES (45% in control schools and 51% in treatment schools).

**Figure 4.     Student SES distribution for Scaling Up Early Reading Intervention—Uttarakhand, by treatment group**
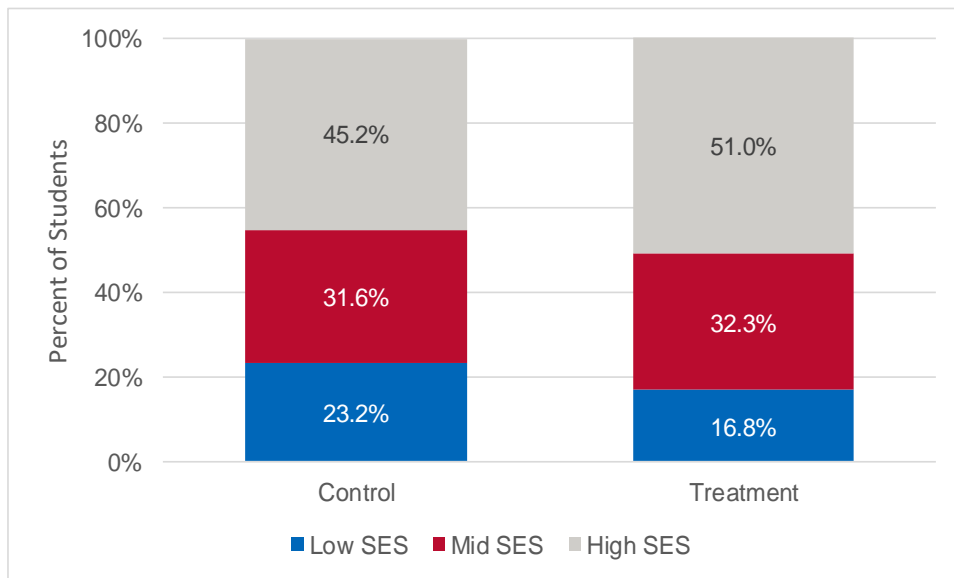


More students in treatment schools reported having help with homework and having extra reading material at home. Across both groups of sampled students, absenteeism was high, as over half of students reported being absent in the previous week (see *Table 11*).

**Table 11.      Student demographics for Scaling Up Early Reading Intervention Project—Uttarakhand, by treatment group**

| | Control<br>Mean (Standard Error) | Treatment<br>Mean (Standard Error) |
|---|---|---|
| Breakfast before school | 90.8% (1.8) | 91.2% (1.4) |
| Help with homework at home | 65.9% (3.9) | 71.5% (2) |
| Extra reading material at home | 44.8% (3.1) | 56.6% (2.5) |
| Absent from school in last week | 59.5% (3.3) | 57.8% (2.1) |
| Someone reads aloud at home | 63.3% (3.1) | 63.9% (1.7) |

### 6.1.2   ASER Results

**Figure 5** below shows performance of Standard 2 students in treatment and control schools for the Scaling Up Early Reading Intervention project in Uttarakhand. Of students in treatment schools, 20.3% could read the Standard 2-level text, and another 23.9% of students could read the Standard 1-level text, but not the Standard 2 text. In control schools, these proportions are 11.1% and 10.4%, respectively. In control schools 28.6% of Standard 2 students could not even read letters, compared to 13.5% of students in treatment schools. Regression analysis (see **Annex I**) confirms that the proportion of students who could read Standard 2-level text in treatment schools was significantly higher than this proportion in control schools, even after controlling for child and household characteristics.

**Figure 5.      Percent of students by ASER reading level for Scaling Up Early Reading Intervention—Uttarakhand, by treatment group**



| | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|
| | Boys | Girls | All children | Boys | Girls | All children |
| ■ Beginner | 31.5 | 26.0 | 28.6 | 14.0 | 13.1 | 13.5 |
| ■ Letter | 37.8 | 45.6 | 41.9 | 36.0 | 30.5 | 33.2 |
| ■ Word | 8.6 | 7.4 | 8.0 | 9.9 | 8.3 | 9.1 |
| ■ Std 1 level text | 12.3 | 8.8 | 10.4 | 22.7 | 25.1 | 23.9 |
| ■ Std 2 level text | 9.9 | 12.2 | 11.1 | 17.5 | 22.9 | 20.3 |
| | Percent of Students | | | | | |
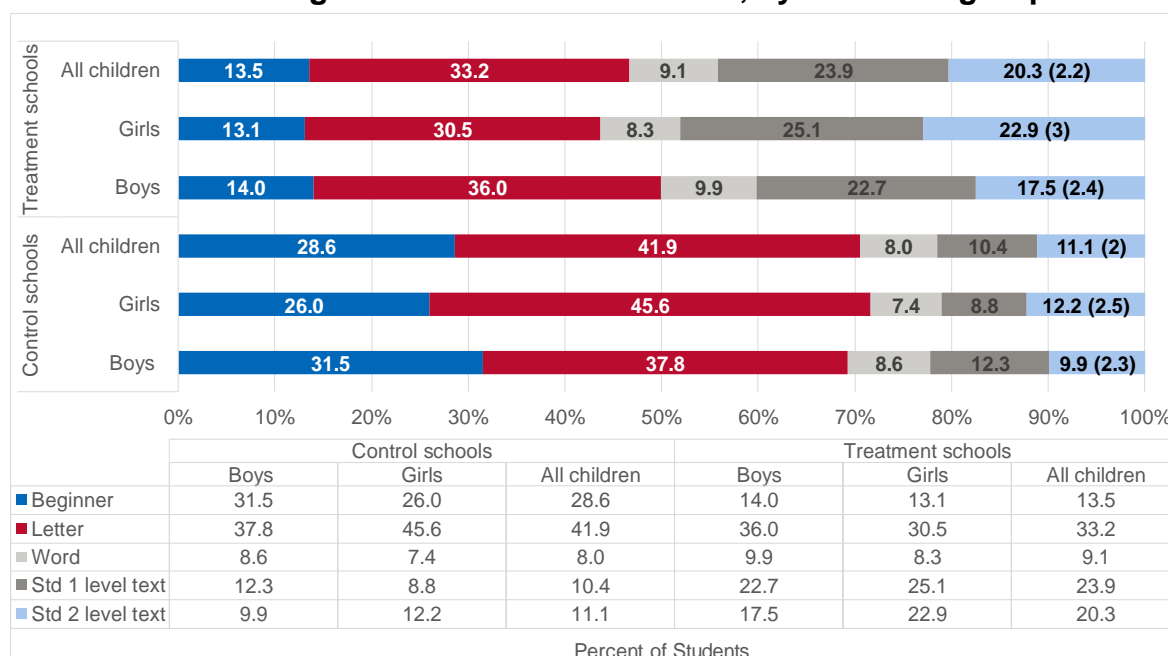
### 6.1.3   EGRA Results

**Table 12** below displays mean scores and effect size in ORF and reading comprehension for Standard 2 students by treatment and control group in the Scaling Up Early Reading Intervention project in Uttarakhand.

**Table 12.    Mean ORF and reading comprehension score by treatment and control group (Scaling Up Early Reading Intervention, Uttarakhand)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 7.3 (1) | 0.49 |
| | Treatment | 15.7 (1.2) | |
| Reading comprehension (% correct) | Control | 7.0% (1.2) | 0.32 |
| | Treatment | 13.7% (1.4) | |

Overall, students in the treatment group outperformed students in the control group. This is partly evident in reading comprehension, where students in the treatment group scored on average approximately 14% correct, compared to 7% correct in the control group. However, the difference is more evident in oral reading—students in the treatment group read 8.4 cwpm more fluently than students in the control group.

Based on an effect size of 0.49 for ORF and 0.32 for reading comprehension, the scores show there was already a small intervention impact conditional to the two groups being balanced (which cannot be assessed).

***Figure 6*** shows the distribution of ORF and reading comprehension scores, respectively, by control and treatment groups.

**Figure 6.    Distribution of ORF, by treatment group (Scaling Up Early Reading Intervention—Uttarakhand)**
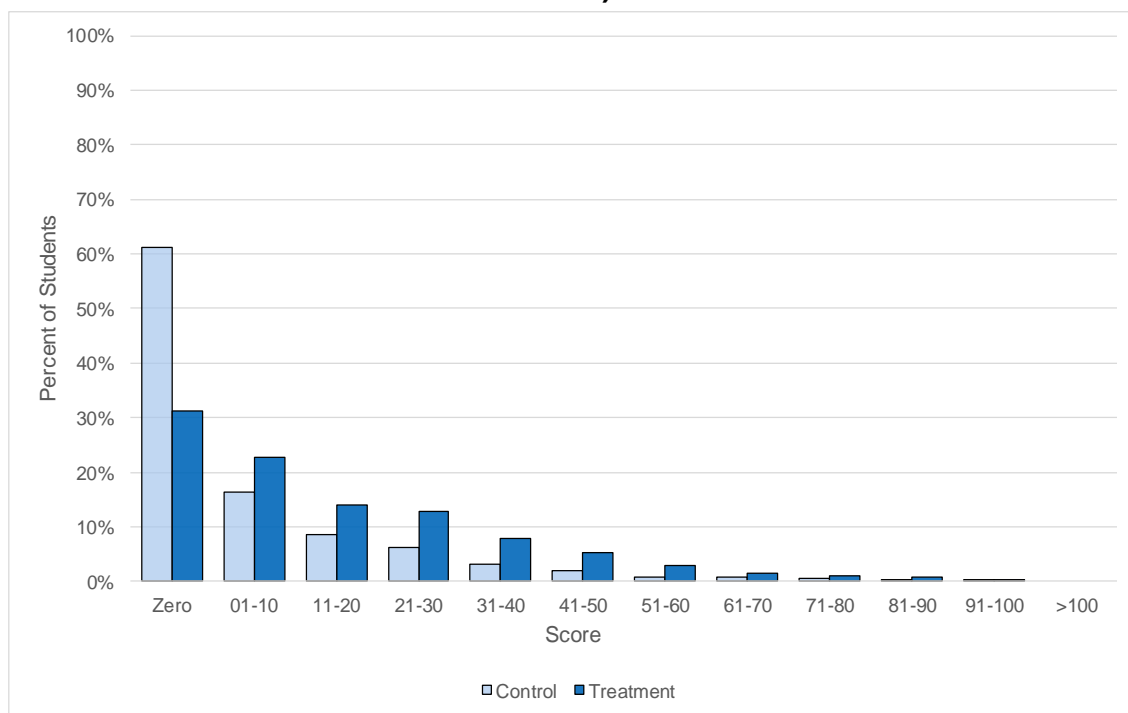


The most evident feature on the graph is the difference across groups in the percentage of students who were unable to read a single word, therefore scoring zero; 61% of students in the control group scored zero, compared to 31% of students in the treatment group. A similar pattern can be seen in ***Figure 7*** for reading comprehension: fewer students in the treatment

group scored zero and instead scored higher in reading comprehension than students in the control group.

**Figure 7.  Distribution of reading comprehension scores, by treatment group (Scaling Up Early Reading Intervention—Uttarakhand)**
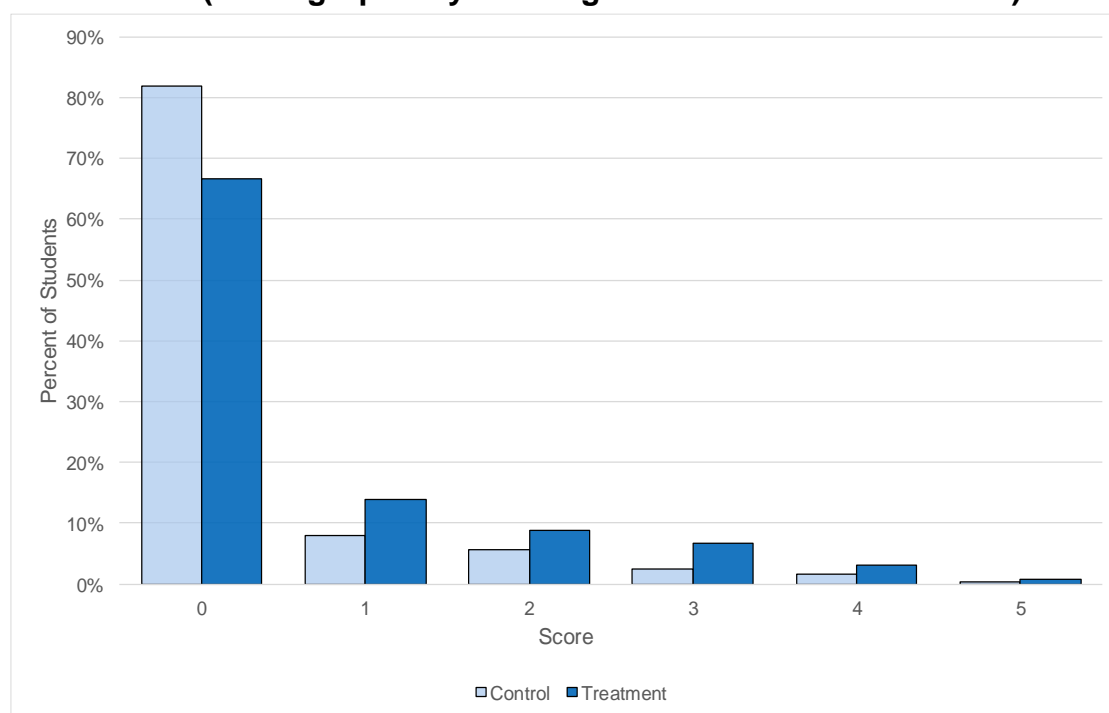


### 6.1.4   ORF per ASER Category

**Table 13** provides the range of ORF scores per each of the ASER reading categories. The majority of students fell within the beginner and letter reading categories and were reading 0 cwpm. Those students who were able to read Standard 2-level text read as few as 5 cwpm with a median of 35 cwpm.

**Table 13.  Range of ORF per ASER reading category (Scaling Up Early Reading Intervention—Uttarakhand)**

| ASER Reading Category | Oral Reading Fluency | | | | | | |
|---|---|---|---|---|---|---|---|
| | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0.08 (0.87) | 23% | 0 | 0 | 0 | 0 | 12 |
| Letter | 1.68 (3.72) | 39% | 0 | 0 | 0 | 2 | 25 |
| Word | 8.77 (7.49) | 8% | 0 | 4 | 8 | 12 | 53 |
| Standard1 | 22.53 (12.11) | 15% | 0 | 13 | 20 | 29 | 70 |
| Standard 2 | 38.3 (19.84) | 14% | 5 | 25 | 35 | 50 | 99 |

### 6.1.5   Scaling Up Early Reading Intervention Project – Chhattisgarh

In Chhattisgarh for the Scaling Up Early Reading Intervention Project, 51% of students assessed in the control group were girls and 49% were boys. In the treatment group, 49% of students were girls and 51% were boys. Most students in Chhattisgarh reported speaking Chhattisgarhi at home (83.8%), which is the same language the assessment was conducted

in. A small number of students indicated they speak Hindi at home (13% in control schools and 17% in treatment schools).

***Figure 8*** shows students from control schools were comparable to students in treatment schools. For both, nearly half of the students were categorized as having low SES; about 30% fell into the mid-SES category, and the remaining 20% were of high SES.

**Figure 8.      Student SES distribution for Scaling Up Early Reading Intervention—Chhattisgarh, by treatment group**
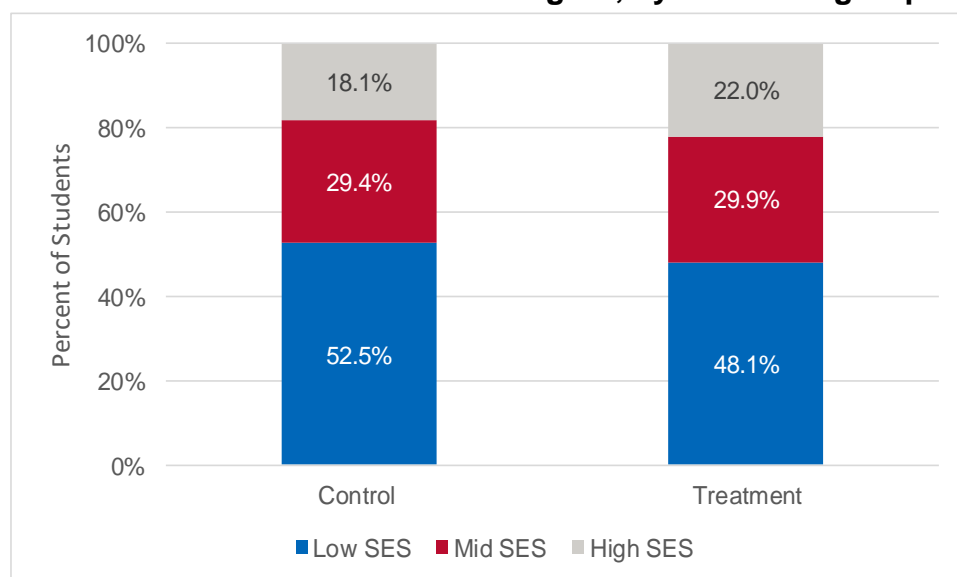


***Table 14*** provides further student demographics and SES variables, again showing little variability between treatment groups. The data presented in ***Table 14*** may imply that fewer students should fall into the lower SES category, as there are large percentages of students who eat breakfast before school and have supplemental support at home. However, it is important to note that the SES distribution is a comparative measure within each project. In order to show variation in the wealth index and split SES among three equal categories, the measure is relative to each project location.

**Table 14.      Student demographics for Scaling Up Early Reading Intervention—Chhattisgarh, by treatment group**

|  | Control Mean (Standard Error) | Treatment Mean (Standard Error) |
|---|---|---|
| Breakfast before school | 96.7% (0.8) | 93.4% (1.4) |
| Help with homework at home | 84.8% (1.6) | 85.1% (1.5) |
| Extra reading material at home | 44.7% (4.2) | 46.8% (4.1) |
| Absent from school in last week | 45.8% (3.7) | 46.9% (2.6) |
| Someone reads aloud at home | 67.6% (2.7) | 67.5% (2.1) |

### *6.1.6   ASER Results*

***Figure 9*** below shows performance of Standard 2 students in treatment and control schools for Scaling Up Early Intervention in Chhattisgarh. Of students in treatment schools, 17% could read Standard 2-level text, while only 6.8% of students in control schools could read Standard 2 level. In control schools 20.1% of Standard 2 students could not read letters,

compared to 11% in treatment schools. Regression analysis (see ***Annex I***) confirms that the proportion of students who could read Standard 2-level text in treatment schools was significantly higher than this proportion in control schools, even after controlling for child and household characteristics.

**Figure 9.     Percent of students by ASER reading level for Scaling Up Early Reading Intervention—Chhattisgarh, by treatment group**



| | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|
| | Boys | Girls | All children | Boys | Girls | All children |
| ■ Beginner | 20.0 | 20.2 | 20.1 | 12.1 | 9.9 | 11.0 |
| ■ Letter | 45.1 | 47.0 | 46.1 | 35.8 | 27.1 | 31.5 |
| ▪ Word | 19.0 | 9.6 | 14.2 | 15.8 | 17.7 | 16.7 |
| ▪ Std 1 level text | 10.9 | 14.8 | 12.9 | 21.8 | 25.7 | 23.7 |
| ▪ Std 2 level text | 5.1 | 8.4 | 6.8 | 14.5 | 19.7 | 17.0 |

Percent of Students
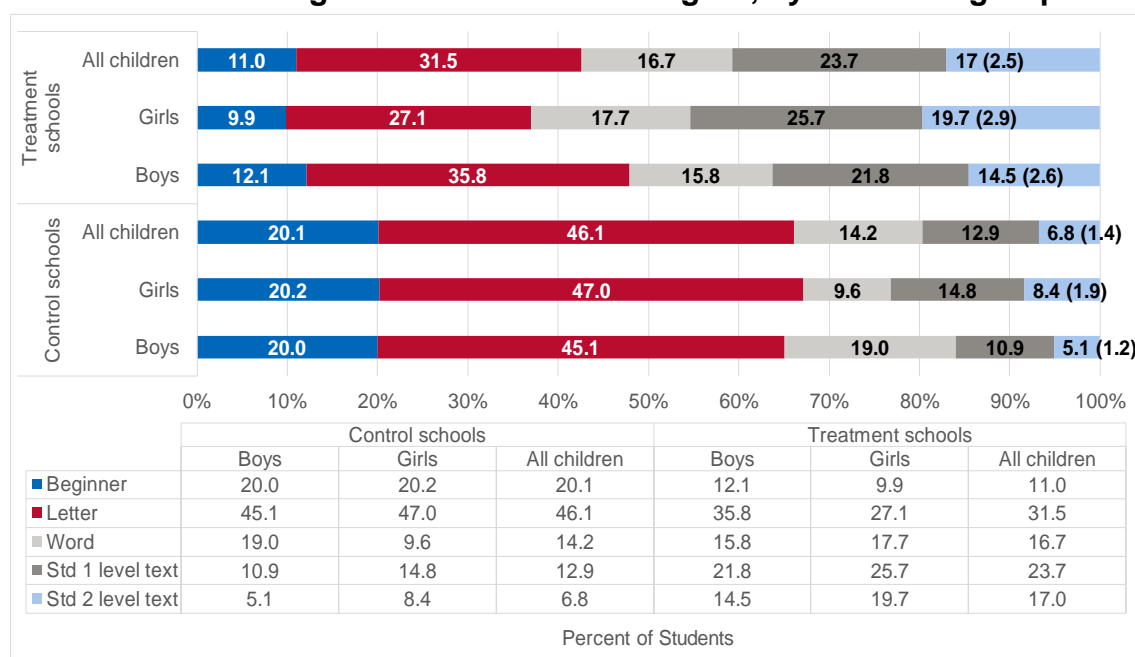
### 6.1.7   EGRA Results

The mean reading scores and effect size in ORF and reading comprehension for Standard 2 students by treatment group in the Scaling Up Early Reading Intervention in Chhattisgarh are shown in ***Table 15***.

**Table 15.     Mean ORF and reading comprehension score by treatment group (Scaling Up Early Reading Intervention—Chhattisgarh)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 6.6 (1) | 0.57 |
| | Treatment | 14.6 (0.9) | |
| Reading comprehension (% correct) | Control | 3.9% (0.8) | 0.38 |
| | Treatment | 9.8% (0.9) | |

Overall, students in the treatment group outperformed students in the control group. This is partly evident in reading comprehension, where students in the treatment group scored on average approximately 10% correct, compared to approximately 4% correct in the control group. However, the difference is more evident in oral reading—students in the treatment group read approximately 8 cwpm more fluently than students in the control group. It would make sense that the magnitude of impact for reading fluency (0.57) is greater than that for comprehension (0.38) as this assessment was done at the start of Standard 2. It would not

be expected that students be reading with comprehension by the end of Standard 1, and comprehension is not as central a focus of the Standard 1 curriculum compared to mastery of phonemes, sounds, and words.

***Figure 10*** below shows the distribution of ORF, by control and treatment groups.

**Figure 10.    Distribution of ORF, by treatment group (Scaling Up Early Reading Intervention—Chhattisgarh)**
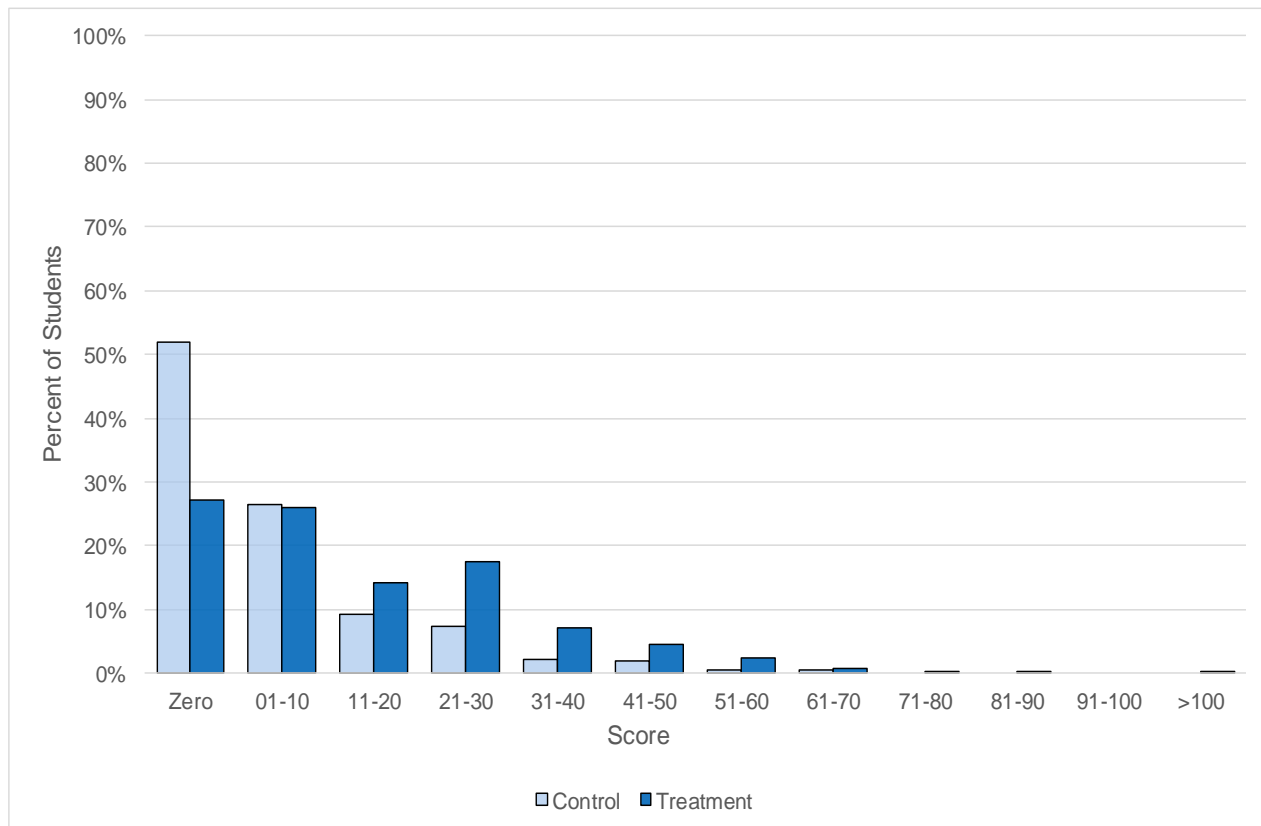


For ORF, the most evident feature on the graph is the difference between the two groups in terms of the percentage of students who were unable to read a single word, therefore scoring zero; over half of students in the control group scored zero, compared to just around a quarter of students in the treatment group. The distribution of scores from the treatment group appears transformed to the right of that of the control group. In other words, the students in the treatment group were moving away from zero scores and towards higher fluencies. A similar pattern can be seen in ***Figure 11*** for reading comprehension: fewer students in the treatment group were scoring zero and were instead scoring higher in reading comprehension than students in the control group. These findings are reflected in the difference in means, discussed above.

**Figure 11.    Distribution of reading comprehension scores, by treatment group (Scaling Up Early Reading Intervention—Chhattisgarh)**
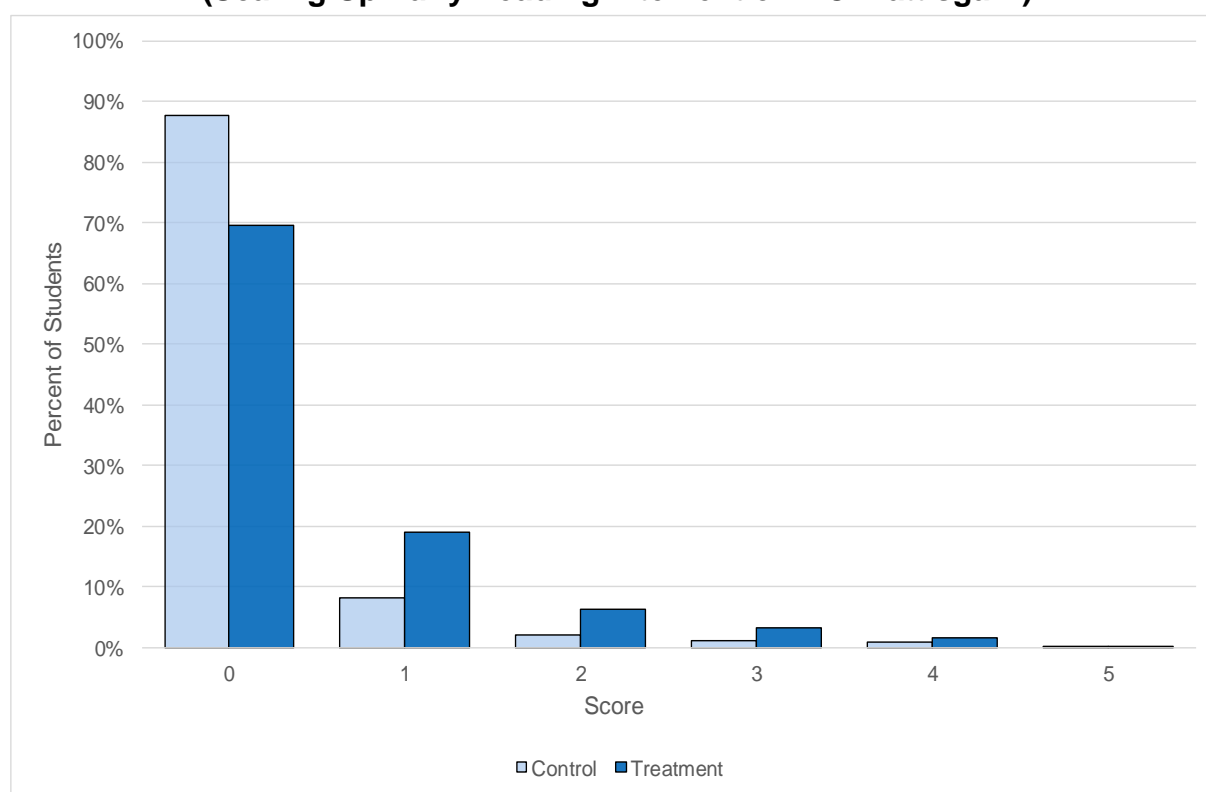


### 6.1.8   ORF per ASER Category

As they did in other project locations, the majority of students in the Scaling Up Early Learning Reading Intervention in Chhattisgarh were categorized at beginner and letter reading levels, reading 0 cwpm. Scores for students categorized at the Standard 2 level ranged from the lowest score of 0 cwpm, to the highest score of 109 cwpm (***Table 16***).

**Table 16.    Range of ORF per ASER reading category (Scaling Up Early Reading Intervention—Chhattisgarh)**

| ASER Reading Category | Oral Reading Fluency | | | | | | |
|---|---|---|---|---|---|---|---|
| | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0.04 (0.38) | 15% | 0 | 0 | 0 | 0 | 5 |
| Letter | 1.57 (3.05) | 39% | 0 | 0 | 0 | 2 | 26 |
| Word | 9.69 (6.08) | 16% | 0 | 5 | 9 | 13 | 33 |
| Standard1 | 22.29 (9.91) | 18% | 1 | 14 | 22 | 28 | 57 |
| Standard 2 | 36.82 (16.04) | 12% | 0 | 25 | 33 | 47 | 109 |

## 6.2    Nurturing Early Literacy

Nurturing Early Literacy is a project co-funded by USAID/India and Tata Trusts, implemented by the Centre for microFinance, 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 have been selected based on an analysis of gaps in education access, delivery, pedagogy, and learning outcomes along with socio-economic 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.

### 6.2.1  *Nurturing Early Literacy – Rajasthan*

For Nurturing Early Literacy in Rajasthan, 48% percent of the students assessed in control schools were girls and 52% were boys. In the treatment group, 49% of students were girls and 51% were boys. While a small number of students responded that they speak Hindi, the language used to administer the assessment, at home (20% in control schools and 9% in treatment schools), the majority of students in both groups indicated that they speak a language other than one of the six language options listed in the assessment (78% in control schools and 89% in treatment schools). The most common "other" language indicated was Grasiya (40.9% of the overall sample).

As shown in **Figure 12**, most sampled students were categorized as having low SES (57% in control schools and 69% in treatment schools); relatively few students fell into the high SES category, especially in treatment schools, where only 5% of students were categorized as having high SES as compared with 17% of students in control schools.

**Figure 12.    Student SES distribution for Nurturing Early Literacy—Rajasthan, by treatment group**
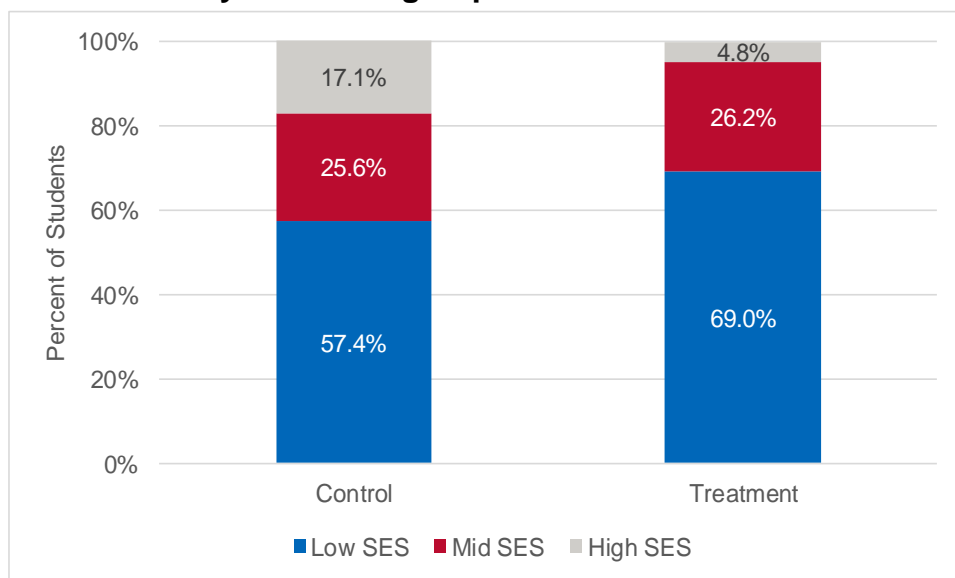


**Table 17** provides further student demographics and SES variables showing little variability between control and treatment groups.

**Table 17.** **Student demographics for Nurturing Early Literacy—Rajasthan, by treatment group**

| | Control Mean & Standard Error | Treatment Mean & Standard Error |
|---|---|---|
| Breakfast before school | 80.3% (2.8) | 81.1% (0.2) |
| Help with homework at home | 42.2% (2.7) | 42.3% (0.4) |
| Extra reading material at home | 32.7% (3.3) | 27.9% (0.3) |
| Absent from school in last week | 54.2% (2.7) | 52.1% (0.4) |
| Someone reads aloud at home | 34.2% (2.6) | 33.0% (0.3) |

### 6.2.2 ASER Results

**Figure 13** below shows the performance of Standard 2 students in treatment and control schools for Nurturing Early Literacy in Rajasthan. Nearly half of the students in both treatment and control schools could not even read basic letters (51% and 53%, respectively). Fewer than 4% of students were able to read the Standard 1-level text or higher in both treatment and control schools. The proportion of students who could read Standard 2-level text in treatment schools was not significantly different from this proportion in control schools (see **Annex I**).

**Figure 13.** **Percent of students by ASER reading level for Nurturing Early Literacy—Rajasthan, by treatment group**



| | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|
| | Boys | Girls | All children | Boys | Girls | All children |
| ■ Beginner | 51.9 | 54.2 | 53.0 | 48.8 | 53.3 | 51.0 |
| ■ Letter | 37.0 | 37.1 | 37.1 | 40.0 | 36.2 | 38.1 |
| ■ Word | 7.3 | 4.8 | 6.1 | 8.8 | 5.6 | 7.2 |
| ■ Std 1 level text | 2.2 | 3.2 | 2.7 | 2.0 | 4.7 | 3.3 |
| ■ Std 2 level text | 1.6 | 0.7 | 1.2 | 0.4 | 0.3 | 0.3 |

Percent of Students
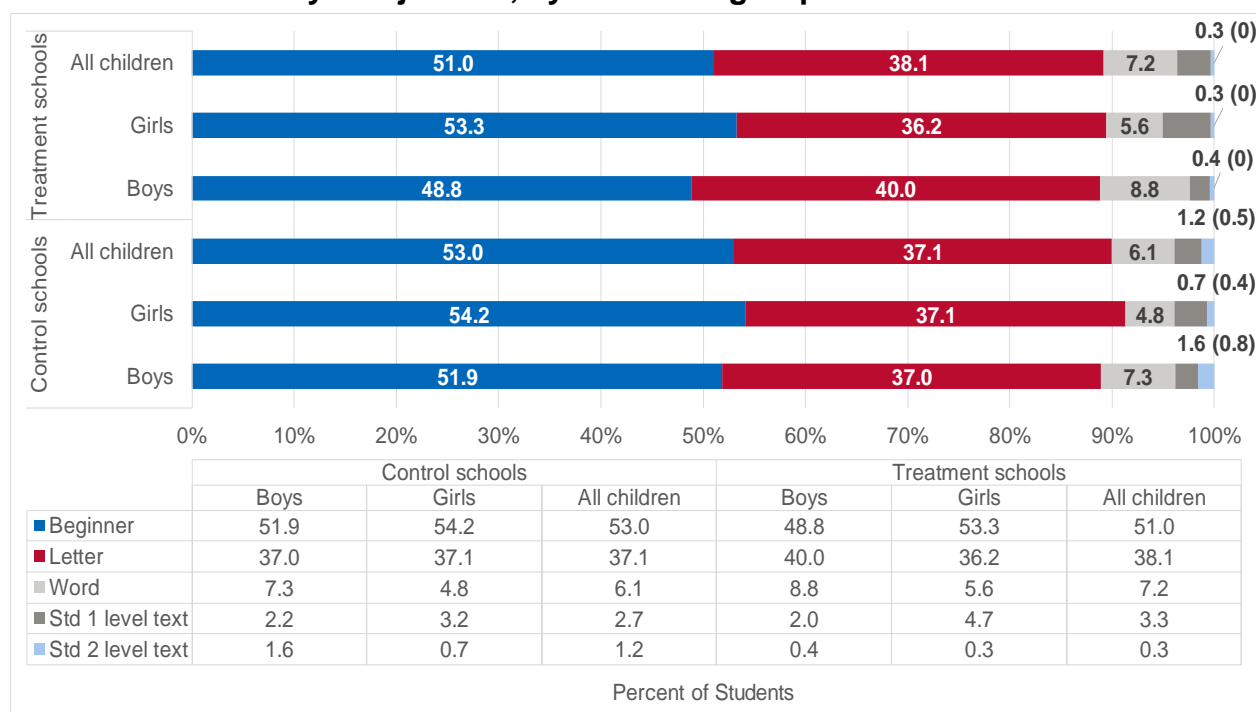
### 6.2.3 EGRA Results

**Table 18** below displays mean scores and effect size in ORF and reading comprehension for Standard 2 students by treatment group in the Nurturing Early Literacy project in Rajasthan.

**Table 18.    Mean ORF and reading comprehension score by treatment group (Nurturing Early Literacy—Rajasthan)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 1.3 (0.3) | 0.06 |
| | Treatment | 1.7 (0) | |
| Reading comprehension (% correct) | Control | 0.7% (0.2) | 0.02 |
| | Treatment | 0.6% (0) | |

Performance was nearly identical across control and treatment groups, with students in control schools reading 1.3 cwpm, and students in treatment schools reading 1.7 cwpm. Because students struggled to read the passage, they also performed poorly in reading comprehension; roughly 97% of students could not answer a single reading comprehension question, leading to mean scores of less than 1% for both groups. To some degree, the low performance seen across both treatment and control schools may be explained by the socio-economic and demographic characteristics of the sample. As discussed earlier, 57% of the control group and 69% of the treatment group were categorized as low SES. Less than a third of students in both groups reported having access to extra reading materials at home, and only around a third reported that someone reads to them at home.

*Figure 14* below shows the distribution of ORF by control and treatment groups. The distribution for both groups is heavily skewed to the right, meaning most students were unable to recognize a single word and scored zero. The graph illustrates two minor differences between the control and treatment groups. The first difference is in zero scores: slightly more students in the control group (approximately 87%) were unable to read a single word compared to students in the treatment group (approximately 81%). It also highlights that a higher percentage of students in the treatment schools performed better than students in the control schools in the 1–10 cwpm range, 13% to 8%, respectively.

**Figure 14.** **Distribution of ORF, by treatment group (Nurturing Early Literacy—Rajasthan)**
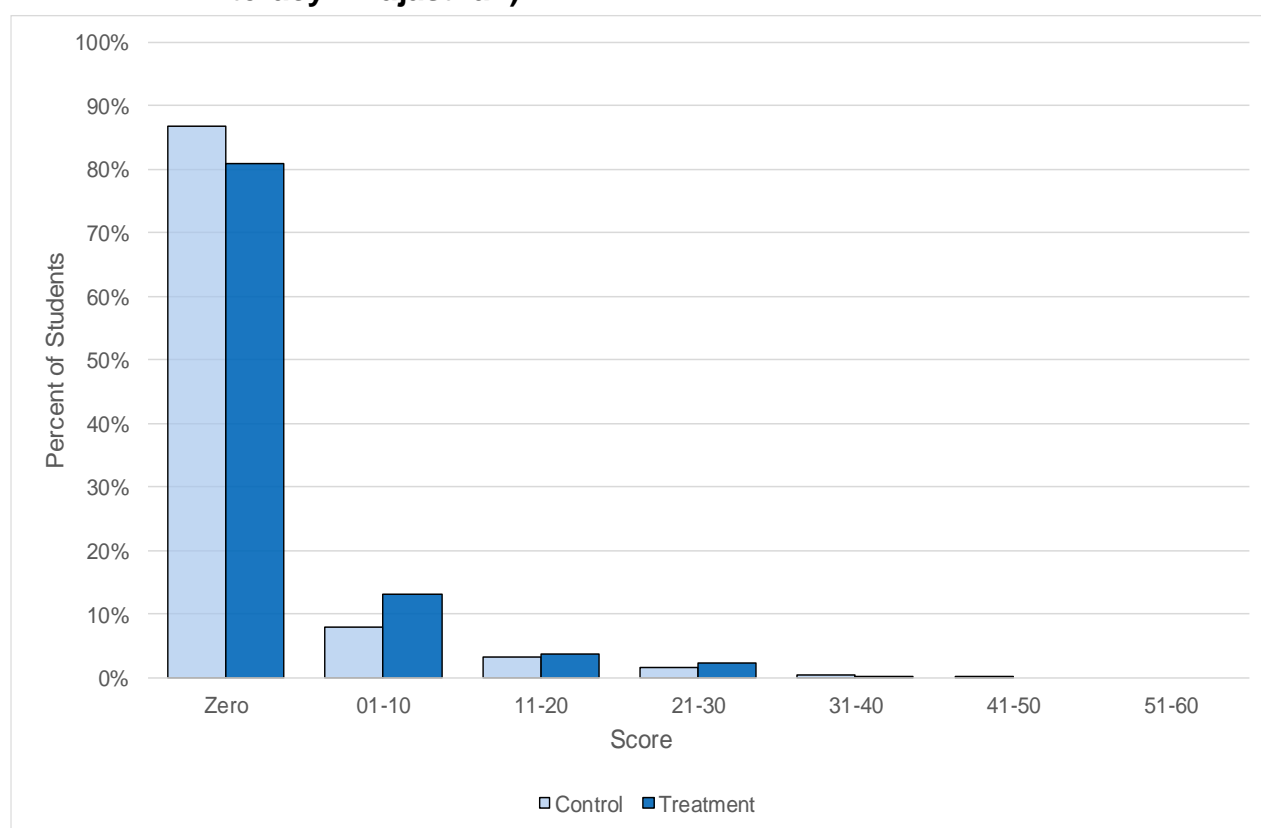


### 6.2.4  ORF per ASER Category

As previously discussed, very few students were able to read Standard 1 or Standard 2 grade-level texts. As presented in *Table 19*, those students who were able to read connected text were reading no more than 41 cwpm. The median score for students reading Standard 1-level text was only 16 cwpm, and only 22 cwpm for Standard 2-level.

**Table 19.** **Range of ORF per ASER Reading Category (Nurturing Early Literacy—Rajasthan)**

| ASER Reading Category | Oral Reading Fluency | | | | | | |
|---|---|---|---|---|---|---|---|
| | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0 (0.03) | 52% | 0 | 0 | 0 | 0 | 1 |
| Letter | 0.64 (1.76) | 37% | 0 | 0 | 0 | 0 | 17 |
| Word | 8.17 (6.86) | 7% | 0 | 3 | 8 | 12 | 34 |
| Standard1 | 17.07 (8.34) | 3% | 0 | 11 | 16 | 23 | 34 |
| Standard 2 | 23.18 (8.82) | 1% | 11 | 16 | 22 | 30 | 41 |

### 6.2.5  Nurturing Early Literacy – Karnataka

Of the students assessed in Karnataka for Nurturing Early Literacy, girls and boys were split evenly in control schools. In the treatment group, 51% of students were girls and 49% were boys. Many students indicated that they speak Kannada at home (46% in control schools and 65% in treatment schools), which is the same language the students were assessed in.

Another large proportion of students responded that they speak a language other than the six language options listed in the assessment (45% in control and 30% in treatment).

It is important to note that 19 control schools were selected from two clusters, Arkera (K) and Kandakoor. Selected schools were receiving similar intervention inputs from an organization called Prajayatna, and these schools are not a true control.

As shown in *Figure 15*, nearly half of students in both groups were categorized as mid SES with another large proportion categorized as low SES; when compared with other project locations, Karnataka had relatively few students in the high-SES category.

**Figure 15.    Student SES distribution for Nurturing Early Literacy—Karnataka, by treatment group**
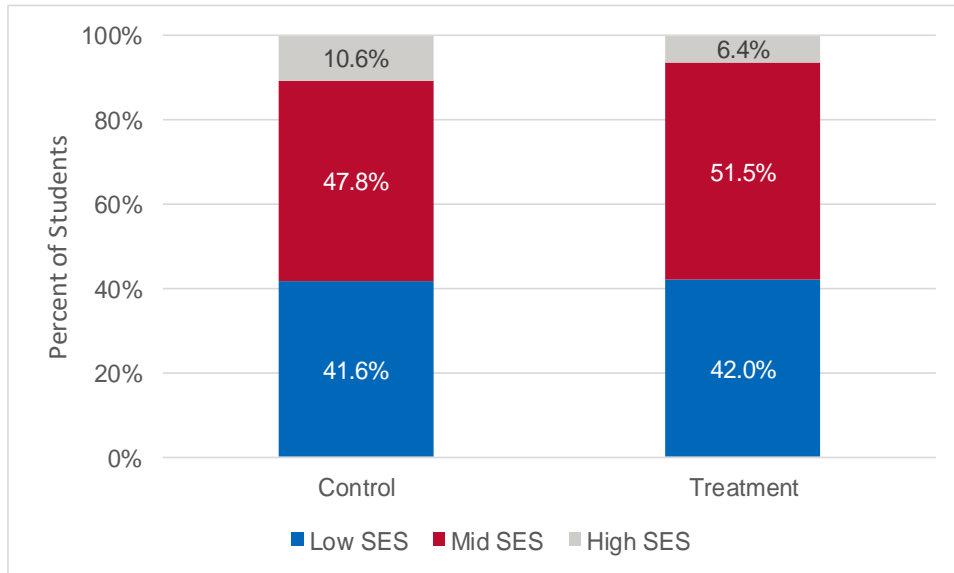


*Table 20* shows students in control and treatment groups were comparable across most variables. Few students in Karnataka reported having extra reading materials at home when compared to students in other project locations.

**Table 20.    Student SES distribution for Nurturing Early Literacy—Karnataka, by treatment group**

|  | Control<br>Mean (Standard Error) | Treatment<br>Mean (Standard Error) |
|---|---|---|
| Breakfast before school | 93.2% (1) | 95.7% (0.7) |
| Help with homework at home | 45.6% (3.1) | 51.6% (2.74) |
| Extra reading material at home | 28.3% (2.3) | 29.8% (2.6) |
| Absent from school in last week | 45.9% (2.8) | 54.1% (2.8) |
| Someone reads aloud at home | 45.6% (2.4) | 42.1% (1.7) |

### 6.2.6   ASER Results

*Figure 16* shows performance of Standard 2 students in treatment and control schools for Nurturing Early Literacy in Karnataka. Nearly half the students in both treatment and control schools were unable to read letters (46.7% in treatment schools and 52.9% in control schools). About 1.5% of Standard 2 students across both treatment and control schools

could read Standard 2-level text. Another 5.4% of students in treatment schools could read Standard 1-level text, but could not read Standard 2-level text. This proportion was 4% in control schools. The proportion of students who could read Standard 2-level text in treatment schools is not significantly different from this proportion in control schools (see **Annex I**).

**Figure 16.    Percent of students by ASER reading level for Nurturing Early Literacy—Karnataka, by treatment group**



| | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|
| | Boys | Girls | All children | Boys | Girls | All children |
| ■ Beginner | 53.1 | 52.7 | 52.9 | 46.7 | 46.7 | 46.7 |
| ■ Letter | 31.4 | 32.5 | 32.0 | 32.3 | 37.8 | 35.1 |
| ■ Word | 10.8 | 8.8 | 9.8 | 12.0 | 10.5 | 11.2 |
| ■ Std 1 level text | 4.0 | 4.0 | 4.0 | 6.5 | 4.3 | 5.4 |
| ■ Std 2 level text | 0.7 | 2.0 | 1.4 | 2.6 | 0.6 | 1.6 |

Percent of Students
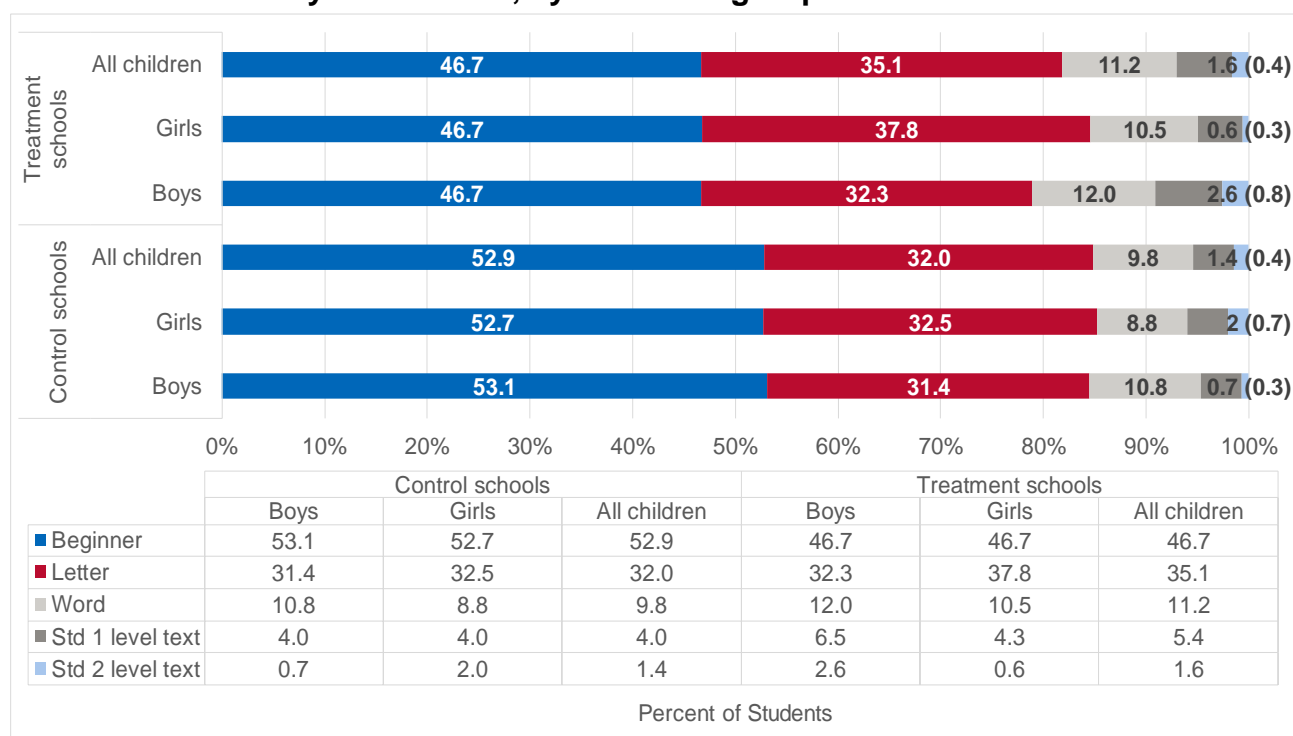
### 6.2.7   EGRA Results

**Table 21** below displays mean scores and effect sizes in ORF and reading comprehension for Standard 2 students by treatment group in Nurturing Early Literacy in Karnataka.

**Table 21.    Mean ORF and reading comprehension score by treatment group (Nurturing Early Literacy—Karnataka)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 2.1 (0.4) | 0.07 |
| | Treatment | 2.6 (0.3) | |
| Reading comprehension (% correct) | Control | 1.0% (0.3) | 0.14 |
| | Treatment | 2.1% (0.4) | |

Performance was nearly identical across control and treatment groups, with students in control schools reading 2.1 cwpm and students in treatment schools reading 2.6 cwpm. Because students struggled to read the passage, they also scored poorly in reading comprehension; roughly 95% of students could not answer a single reading comprehension question, leading to mean scores of less than 2% for both groups. This low performance is supported by the socioeconomic and demographic characteristics of the sample. Less than a

third of students in both groups reported having access to extra reading materials at home, which means student may not be getting the proper support they need outside of school to facilitate learning.

*Figure 17* below shows the distribution of ORF by control and treatment groups. The distribution for both groups is heavily skewed to the right, meaning most students were unable to recognize a single word. This graph illustrates two minor differences between the control and treatment groups. The first difference is in zero scores: slightly more students in the control group (approximately 72%) were unable to read a single word compared to students in the treatment group (approximately 66%). It also highlights that a higher percentage of students in the treatment schools performed better than students in the control schools in the 1–10 cwpm range, 26% to 21%, respectively. This is a typical pattern that is seen in distributions of reading fluency as reading skills develop, but the difference right now is very small.

**Figure 17.    Distribution of ORF, by treatment group (Nurturing Early Literacy—Karnataka)**
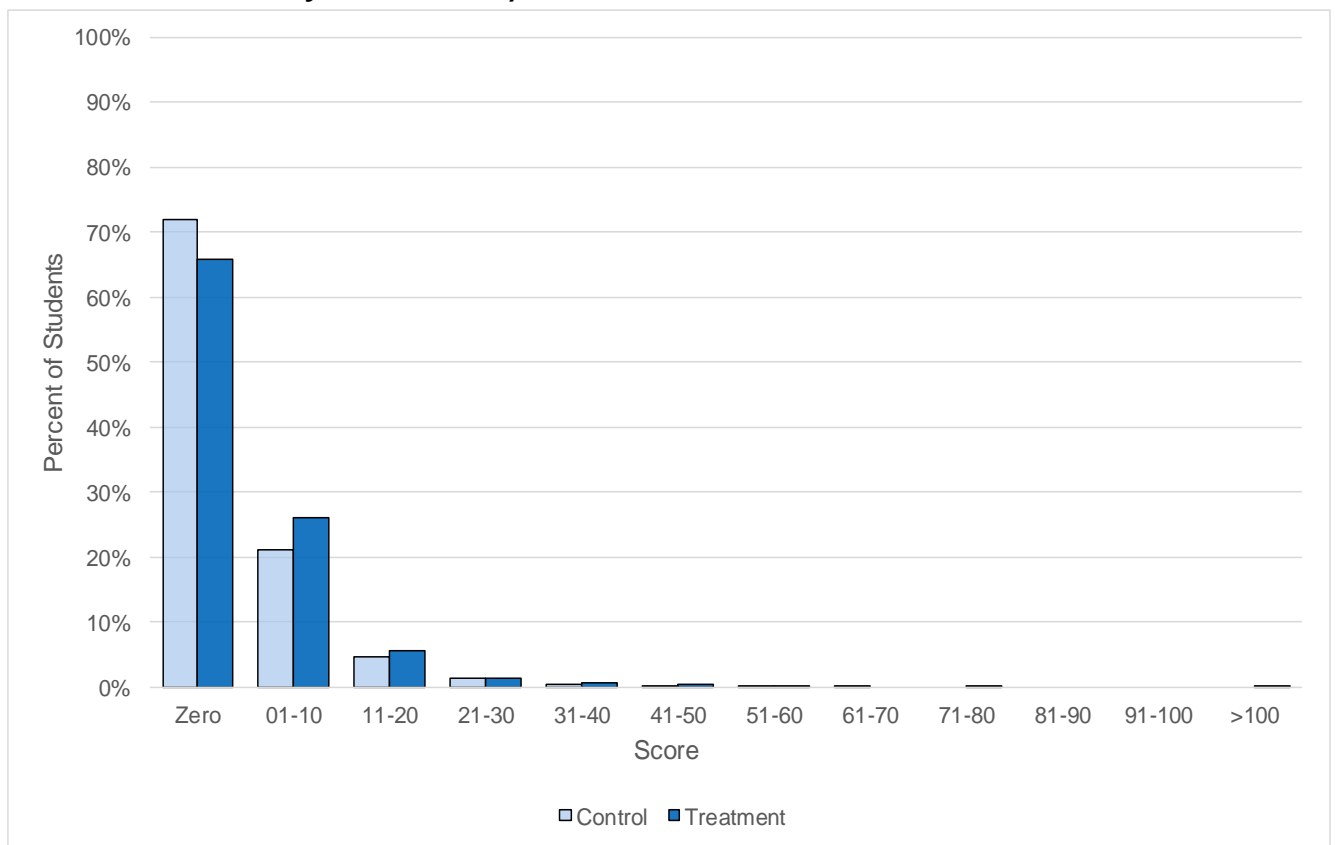


### 6.2.8   ORF per ASER Category

Nearly half of the students in the Nurturing Early Literacy program in Karnataka were categorized at the beginner level, reading 0 cwpm. A small percentage of students who were categorized at the word reading level were reading about 6 cwpm. Among the few students who were able to read Standard 2, scores ranged between 13 cwpm and 150 cwpm (see *Table 22)*.

**Table 22.    Range of ORF per ASER reading category (Nurturing Early Literacy—Karnataka)**

| | Oral Reading Fluency |
|---|---|

| ASER Reading Category | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
|---|---|---|---|---|---|---|---|
| Beginner | 0.13 (1.09) | 49% | 0 | 0 | 0 | 0 | 18 |
| Letter | 1.13 (2.07) | 34% | 0 | 0 | 0 | 2 | 16 |
| Word | 6.66 (5) | 11% | 0 | 3 | 6 | 10 | 25 |
| Standard1 | 15.39 (9.11) | 5% | 0 | 11 | 13 | 20 | 48 |
| Standard 2 | 38.24 (27.38) | 2% | 13 | 22 | 34 | 46 | 150 |

### 6.2.9 *Nurturing Early Literacy – Maharashtra*

In Maharashtra for the Nurturing Early Literacy project, 47% percent of the students assessed in the control group were girls and 53% were boys. In the treatment group, 51% of student were girls and 49% were boys. Nearly 97% of all sampled students in Maharashtra speak Marathi, which is also the language the students were assessed in.

The largest proportion of students in both control and treatment schools were categorized as having high SES (see *Figure 18*); the percent of students with high SES was slightly higher in treatment schools (82%) than control schools (73%). Overall, this project location had the highest number of students categorized as having high SES compared to other project locations. Very few students were categorized as having low SES (about 2%). Interestingly, SES student distributions for Nurturing Early Literacy were significantly different than for sampled students for the Right to Read project also being implemented in Maharashtra.

**Figure 18.    Student SES distribution for Nurturing Early Literacy— Maharashtra, by treatment group**
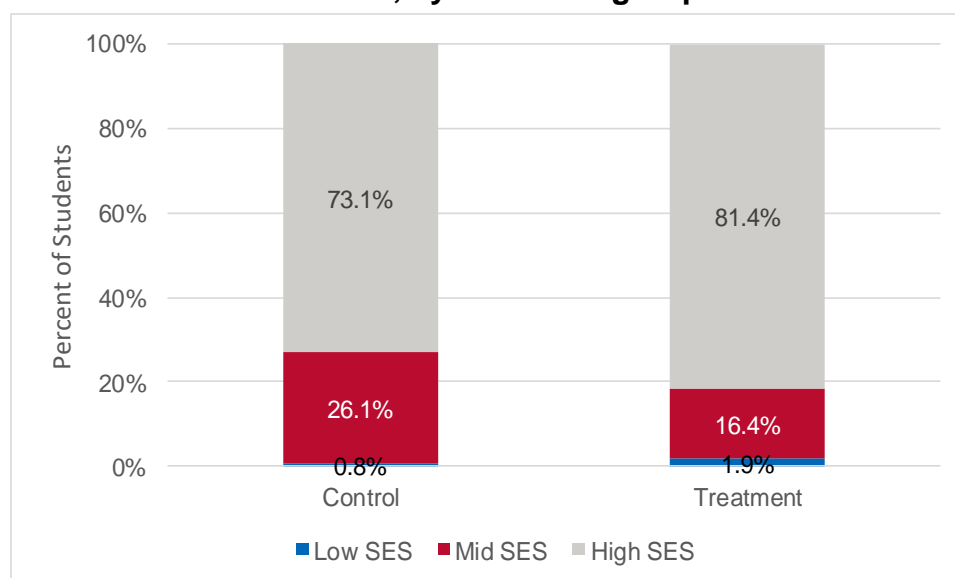


As shown in *Table 23,* students from treatment and control schools were comparable across the various demographic and SES variables. The percent of students who indicated having access to reading material at home was much higher in this project location (55% in control and 65% in treatment schools) compared to other project locations. These schools also had more students reporting that they received help with homework at home (78% in control schools and 85% in treatment schools). Relatively few students in both treatment and control indicated being absent from school in the previous week (about 37%).

**Table 23.    Student demographics for Nurturing Early Literacy—Maharashtra, by treatment group**

|  | Control<br>Mean (Standard Error) | Treatment<br>Mean (Standard Error) |
|---|---|---|
| Breakfast before school | 92.9% (1) | 93.7% (0.9) |
| Help with homework at home | 78.2% (1.8) | 85.4% (1.6) |
| Extra reading material at home | 55.6% (2.4) | 65% (2.4) |
| Absent from school in last week | 37.9% (3.6) | 36.1% (2.8) |
| Someone reads aloud at home | 73.0% (2.1) | 76.8% (2) |

### 6.2.10 ASER Results

***Figure 19*** below shows performance of Standard 2 students on the ASER reading assessment in treatment and control schools for Nurturing Early Literacy in Maharashtra. Of students in treatment schools, 53.9% could read Standard 2-level text and another 25.7% students could read Standard 1-level text, but not Standard 2-level text. Compared to all other programs, student performance in schools for this project location was fairly high. This can be attributed to Satara's (the district in Maharashtra where this project is based) higher than average learning levels.[7] Only 1.9% of students in treatment schools were unable to read letters.

In control schools, 57.1% students could read Standard 2-level text and 27.2% students could read Standard 1-level text, but not Standard 2-level text. The proportion of students who could read Standard 2-level text in treatment schools is not significantly different from that in control schools (see ***Annex I***).

---

[7] See district-level estimates for Maharashtra for ASER 2014: (footnote continued on the following page) http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202014/District%20Estimates/maharashtra.pdf and ASER 2016: http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/District%20Estimates/maharashtra.pdf

**Figure 19.    Percent of students by ASER reading level for Nurturing Early Literacy—Maharashtra, by treatment group**



| | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|
| | Boys | Girls | All children | Boys | Girls | All children |
| ■ Beginner | 2.7 | 1.8 | 2.3 | 2.4 | 1.4 | 1.9 |
| ■ Letter | 6.6 | 3.7 | 5.2 | 9.4 | 5.0 | 7.2 |
| ■ Word | 10.2 | 6.0 | 8.2 | 13.6 | 9.0 | 11.3 |
| ■ Std 1 level text | 31.1 | 22.8 | 27.2 | 27.0 | 24.3 | 25.7 |
| ■ Std 2 level text | 49.5 | 65.7 | 57.1 | 47.6 | 60.3 | 53.9 |

Percent of Students
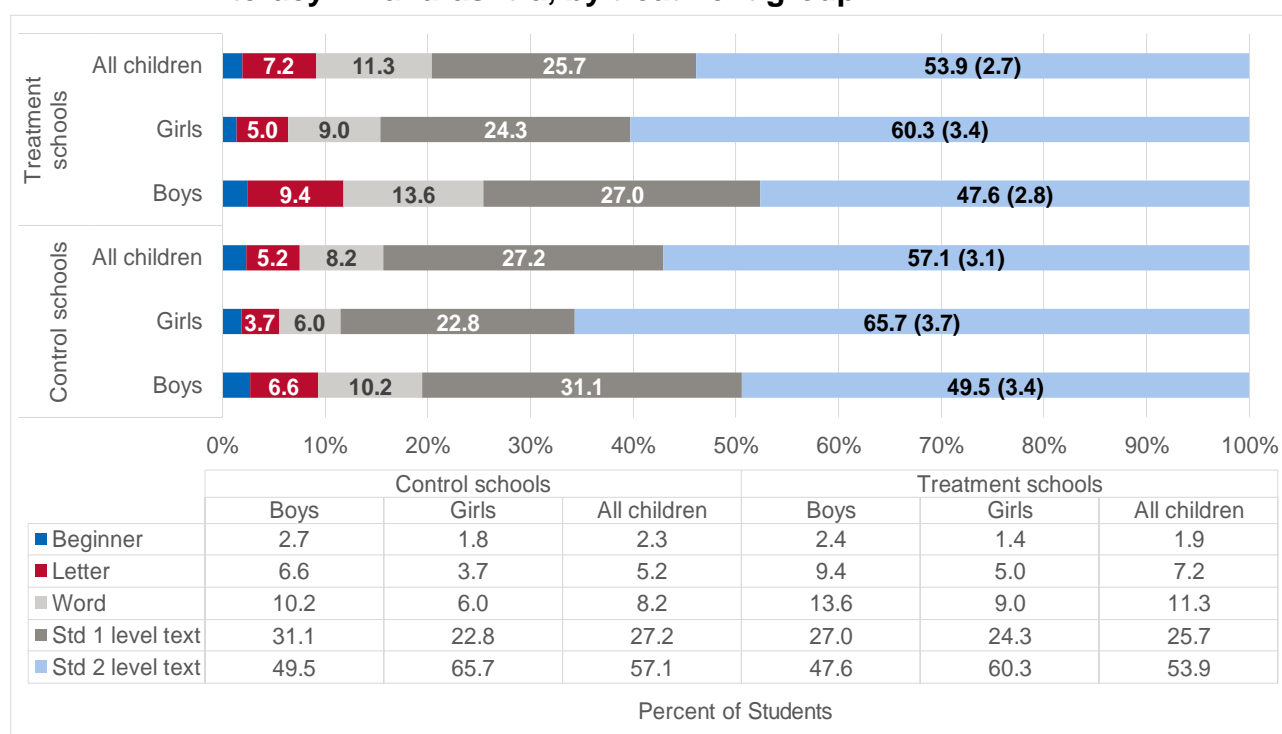
### *6.2.11 EGRA Results*

**Table 24** below displays mean scores and effect sizes in ORF and reading comprehension for Standard 2 students by treatment group in the Nurturing Early Literacy project in Maharashtra.

**Table 24.    Mean ORF and reading comprehension score by treatment group (Nurturing Early Literacy—Maharashtra)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 30.2 (0.9) | 0.29 |
| | Treatment | 35.3 (1.1) | |
| Reading comprehension (% correct) | Control | 40.5% (1.3) | 0.12 |
| | Treatment | 44.0% (2) | |

Overall, students in the Nurturing Early Reading project in Maharashtra demonstrated the highest reading proficiency across all project locations. As this initial assessment is being conducted towards the end of the project lifecycle, the treatment scores are likely a combination of the project's impact along with other socioeconomic factors; most students in Maharashtra fell into the highest socioeconomic group—73% of students in control schools and 82% of students in treatment schools. While the performance in reading comprehension was slightly more matched between the two groups compared to other project locations, students in the treatment group still significantly outperformed students in the control group by approximately 5 cwpm in the oral reading subtask.

A small effect size of 0.29 for ORF indicates that students scored better in treatment than in control schools. However, it is not possible to fully attribute this difference to the intervention, as more students in this project location came from the highest socioeconomic group.

*Figure 20* and *Figure 21* show the distribution of ORF and reading comprehension scores, respectively, by control and treatment groups. For ORF, the distribution for both groups appears to be a similar shape, but more students from treatment schools were to the right of the center than in the control group, confirming the results seen in *Table 24.*

**Figure 20.    Distribution of ORF, by treatment group (Nurturing Early Literacy—Maharashtra)**
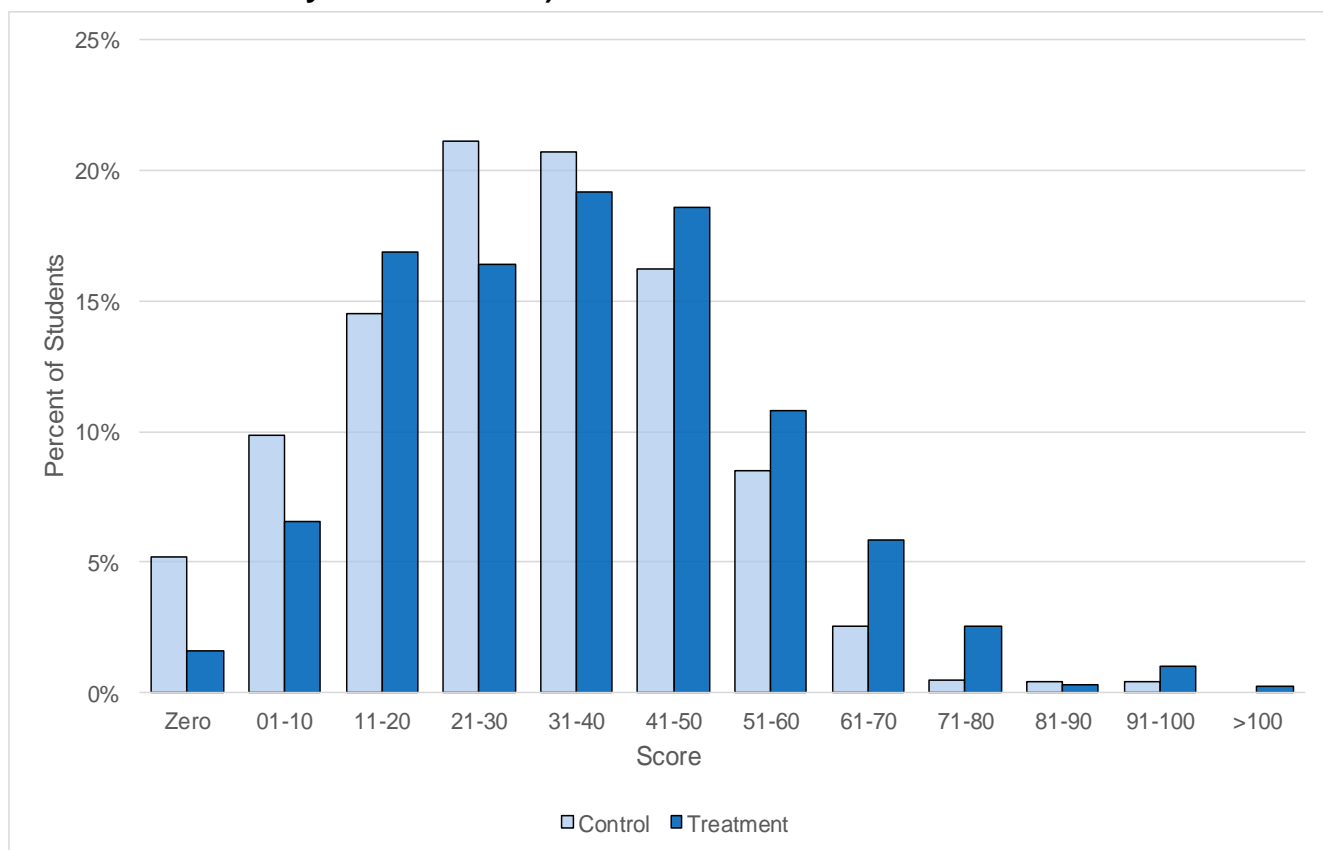


A similar pattern is highlighted in *Figure 21*, reading comprehension scores by treatment and control group, where more students in the treatment group scored 4 or 5.

**Figure 21.    Distribution of reading comprehension scores, by treatment group (Nurturing Early Literacy—Maharashtra)**
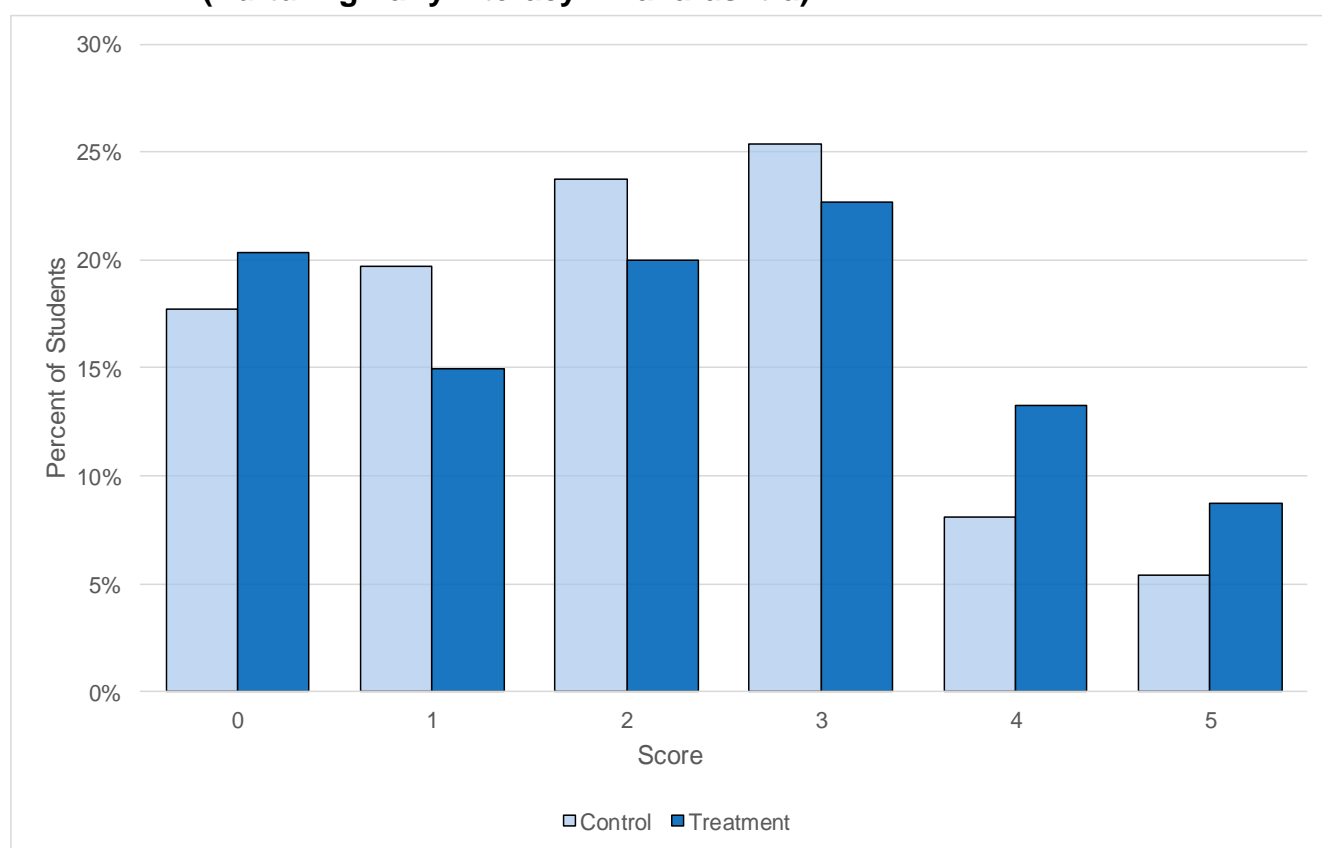


### 6.2.12 ORF per ASER Category

As discussed earlier, the majority of students in this project location were reading Standard 2-level text. **Table 25** shows that students categorized at this level had a vast range of reading abilities, reading between 0 and 116 correct words with a median ORF of 42 cwpm. The median score for students reading at Standard 2 level is slightly higher than the mean ORFs in control and treatment groups, 30 and 35 cwpm, respectively.

**Table 25.    Range of ORF by ASER reading categories (Nurturing Early Literacy—Maharashtra)**

| ASER Reading Category | Oral Reading Fluency | | | | | | |
|---|---|---|---|---|---|---|---|
| | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0 (0) | 2% | 0 | 0 | 0 | 0 | 0 |
| Letter | 4.82 (5.38) | 6% | 0 | 0 | 4 | 9 | 30 |
| Word | 12.81 (7.9) | 10% | 0 | 9 | 12 | 17 | 49 |
| Standard1 | 25.6 (11.06) | 26% | 2 | 18 | 24 | 32 | 96 |
| Standard 2 | 43.05 (16.32) | 55% | 0 | 32 | 42 | 52 | 116 |

## 6.3    Teacher Innovations in Practice – Uttar Pradesh

The Teacher Innovations in Practice project is implemented by Schools and Teachers Innovating for Results (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 four-year project will conclude in September 2018. The project aims to reach 564,000 learners in Standards 1–5. For the assessment, the project assessed learners in Hindi from 10 districts in Uttar Pradesh.[8]

For Teacher Innovations in Practice in Uttar Pradesh, 54% percent of the students assessed in the control group were girls and 46% were boys. In the treatment group, an equal number of boys and girls were assessed. A large majority of students in both control and treatment schools responded that they speak Hindi at home, which was also the language of the assessment (70% in control schools and 84% in treatment schools). A small percentage of students (5% in control schools and 3% in treatment schools) indicated that they speak a language other than one of the six language options listed in the assessment.

*Figure 22* shows that SES levels were comparable across control and treatment schools. The distribution of SES levels among sampled students for this project was comparable to that of student SES levels in the Start Early: Read in Time project, also being implemented in schools in Uttar Pradesh. Students in control and treatment groups are split evenly across low, mid, and high SES, with nearly one-third of students accounting for each SES level.

**Figure 22.    Student SES distribution for Teacher Innovations in Practice—Uttar Pradesh, by treatment group**
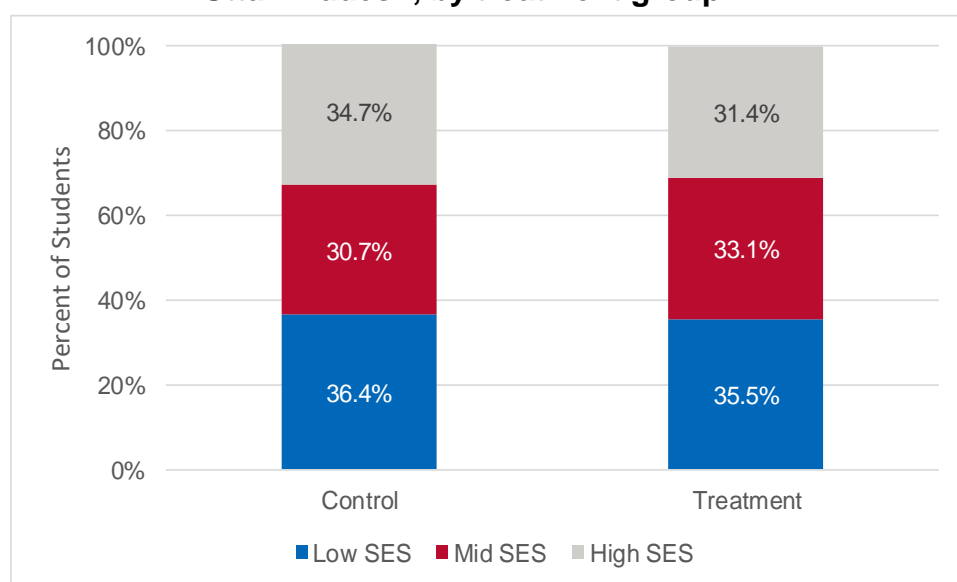


As shown in *Table 26*, slightly higher numbers of students from control schools indicated that they had help with homework at home, extra reading materials at home, and someone who read aloud at home, and that they were absent from school in the previous week. In both groups, the percentage of students who indicated having extra reading materials at home was relatively low (about 30% across both groups).

---

[8] Implementation rollout at the district level as follows: one district in 2014 (Lucknow), 3 districts in 2015 (Raebareli, Varanasi, Unnao), 3 districts in 2016 (Faizabad, Kanpur, Jaunpur), and 3 districts in 2017 (Barabanki, Chandoli, Mirzapur).

**Table 26.    Student demographics for Teacher Innovations in Practice—Uttar Pradesh, by treatment group**

|  | Control Mean (Standard Error) | Treatment Mean (Standard Error) |
|---|---|---|
| Breakfast before school | 88.9% (1.9) | 91.0% (1.6) |
| Help with homework at home | 61.2% (3.7) | 59.4% (3.1) |
| Extra reading material at home | 34.1% (4) | 25.6% (2.7) |
| Absent from school in last week | 54.4% (2.5) | 52.8% (4) |
| Someone reads aloud at home | 48.7% (3.4) | 40.1% (2.7) |

### 6.3.1   ASER Results

**Figure 23** below shows performance of Standard 2 students in treatment and control schools for Teacher Innovations in Practice in Uttar Pradesh. Of students in treatment schools, 6.1% were able to read Standard 2-level text, and another 2.7% were able to read Standard 1-level text, but not Standard 2-level text. The percentages for control schools were 4.9% and 2.9%, respectively. The percent of students who could read Standard 2-level text in treatment schools was not significantly different from that in control schools (see **Annex I**).

**Figure 23.    Percent of students by ASER reading level for Teacher Innovations in Practice—Uttar Pradesh, by treatment group**



|  |  | Control schools |  |  | Treatment schools |  |  |
|---|---|---|---|---|---|---|---|
|  |  | Boys | Girls | All children | Boys | Girls | All children |
| ■ | Beginner | 44.0 | 39.3 | 41.5 | 41.8 | 46.5 | 44.1 |
| ■ | Letter | 42.5 | 44.3 | 43.5 | 41.6 | 42.6 | 42.1 |
| ■ | Word | 7.1 | 7.4 | 7.2 | 6.7 | 3.3 | 5.0 |
| ■ | Std 1 level text | 2.8 | 3.0 | 2.9 | 3.3 | 2.2 | 2.7 |
| ■ | Std 2 level text | 3.7 | 6.0 | 4.9 | 6.7 | 5.5 | 6.1 |
|  |  |  | Percent of Students |  |  |  |  |

### 6.3.2   EGRA Results

**Table 27** displays mean scores and effect size in ORF and reading comprehension for Standard 2 students by treatment group in the STIR project in Uttar Pradesh.

**Table 27.    Mean ORF and reading comprehension score by treatment group (Teacher Innovations in Practice—Uttar Pradesh)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 3.1 (0.6) | 0.06 |
| | Treatment | 3.7 (0.6) | |
| Reading comprehension (% correct) | Control | 3.4% (0.8) | 0.01 |
| | Treatment | 3.3% (0.6) | |

Performance was similar across control and treatment groups, with students in control schools reading 3.1 cwpm and students in treatment schools reading 3.7 cwpm. Overall, student performance in reading comprehension was low; roughly 92% of students could not answer a single reading comprehension question, which was also reflected in their mean scores of less than 4% for both groups.

*Figure 24* below shows the distribution of ORF by control and treatment groups. The distribution for both groups is heavily skewed to the right, with the majority of students unable to recognize a single word and scoring zero. The figure below reveals nearly identical distributions among the two groups.

**Figure 24.    Distribution of ORF, by treatment group (Teacher Innovations in Practice—Uttar Pradesh)**
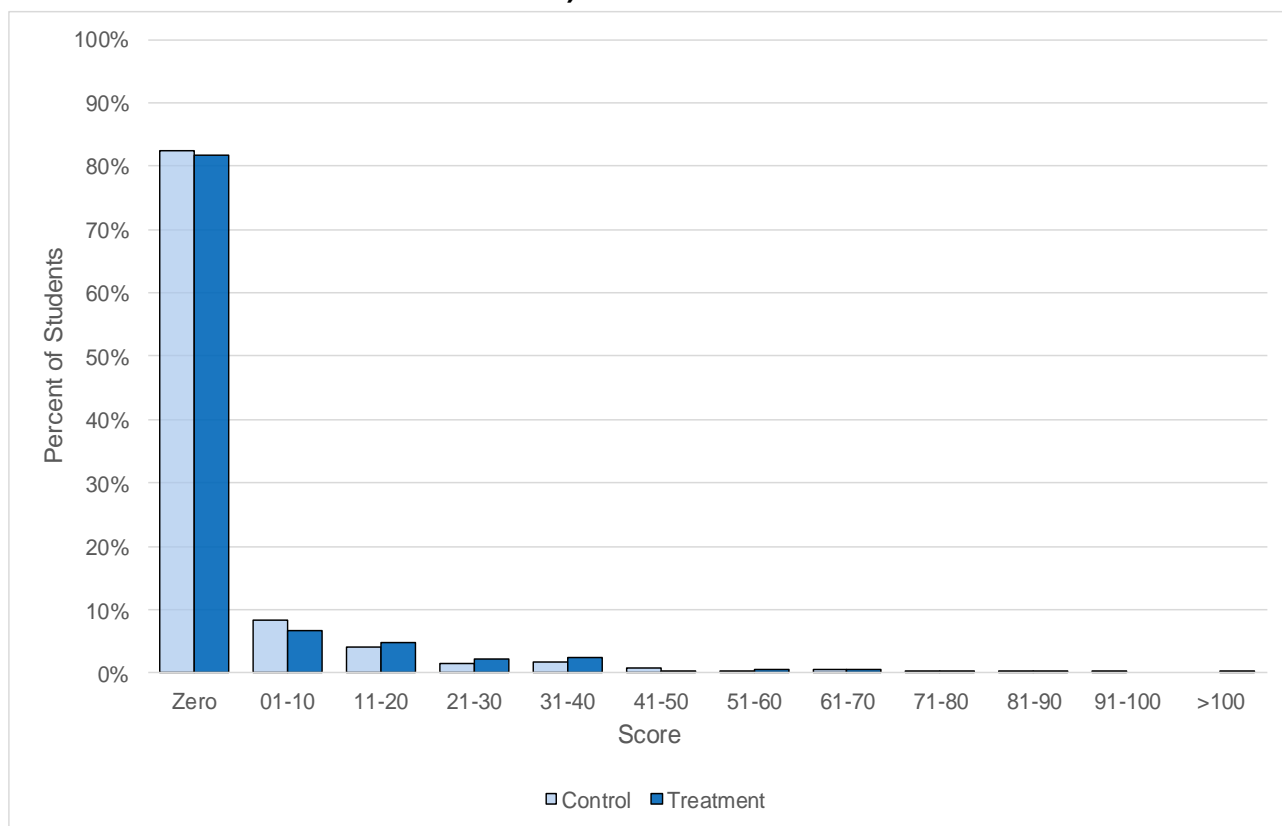


### 6.3.3    ORF per ASER Category

As shown in *Table 28*, 85% of students in this project location were classified at the beginner and letter reading levels, reading 0 cwpm. Students who were able to read

Standard 2-level text had had ORF scores ranging between 0 and 105 cwpm, with a median of 37 cwpm.

**Table 28.    Range of ORF per ASER reading category (Teacher Innovations in Practice—Uttar Pradesh)**

| ASER Reading Category | Oral Reading Fluency | | | | | | |
|---|---|---|---|---|---|---|---|
| | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0.03 (0.44) | 42% | 0 | 0 | 0 | 0 | 9 |
| Letter | 0.75 (2.31) | 43% | 0 | 0 | 0 | 0 | 24 |
| Word | 5.85 (7.3) | 7% | 0 | 0 | 4 | 12 | 33 |
| Standard1 | 17.75 (7.09) | 3% | 0 | 14 | 17 | 22 | 36 |
| Standard 2 | 38.15 (19.8) | 5% | 0 | 24 | 37 | 47 | 105 |

## 6.4    Start Early: Read in Time

CARE India: India Solutions for Sustainable Development started implementing Start Early: Read in Time in July 2014. The project is currently in the fourth and final year of implementation. The project focuses on enhancing reading skills of Standard 1–4 students in formal primary schools in Uttar Pradesh and Odisha. The aim is to reach 100,000 learners 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.

### 6.4.1    Start Early: Read in Time – Uttar Pradesh

Of the students assessed in Uttar Pradesh for Start Early: Read in Time, 55% in control schools were girls and 45% were boys. In the treatment group, 47% were girls and 53% were boys. The majority of the students in Uttar Pradesh reported speaking Hindi, the language of the assessment, at home (70% in control schools and 84% in treatment schools). Nearly 30% of students in control schools and 16% in treatment schools indicated that they speak a language other than one of the six language options listed in the assessment.

As shown in **Figure 25**, students in treatment and control groups were comparable in terms of SES. Within both groups, SES levels were split evenly across low, mid, and high SES, with nearly 1/3 of students accounting for each SES level.

**Figure 25.    Student SES distribution for Start Early: Read in Time—Uttar Pradesh, by treatment group**
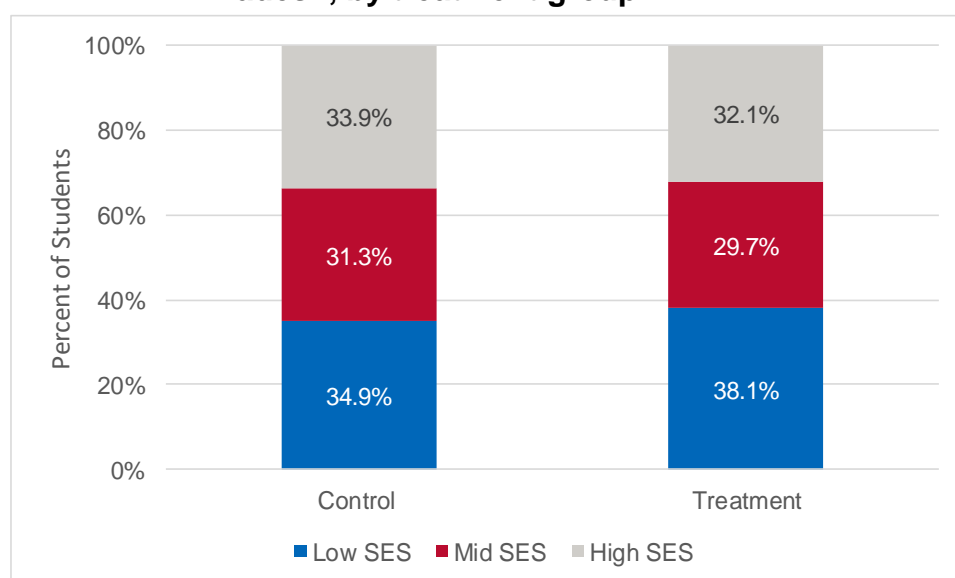


**Table 29** provides further student demographics and SES variables, again showing much similarity between control and treatment groups. The percentages of students in both groups who reported having extra reading material at home was relatively low compared to students in other projects locations.
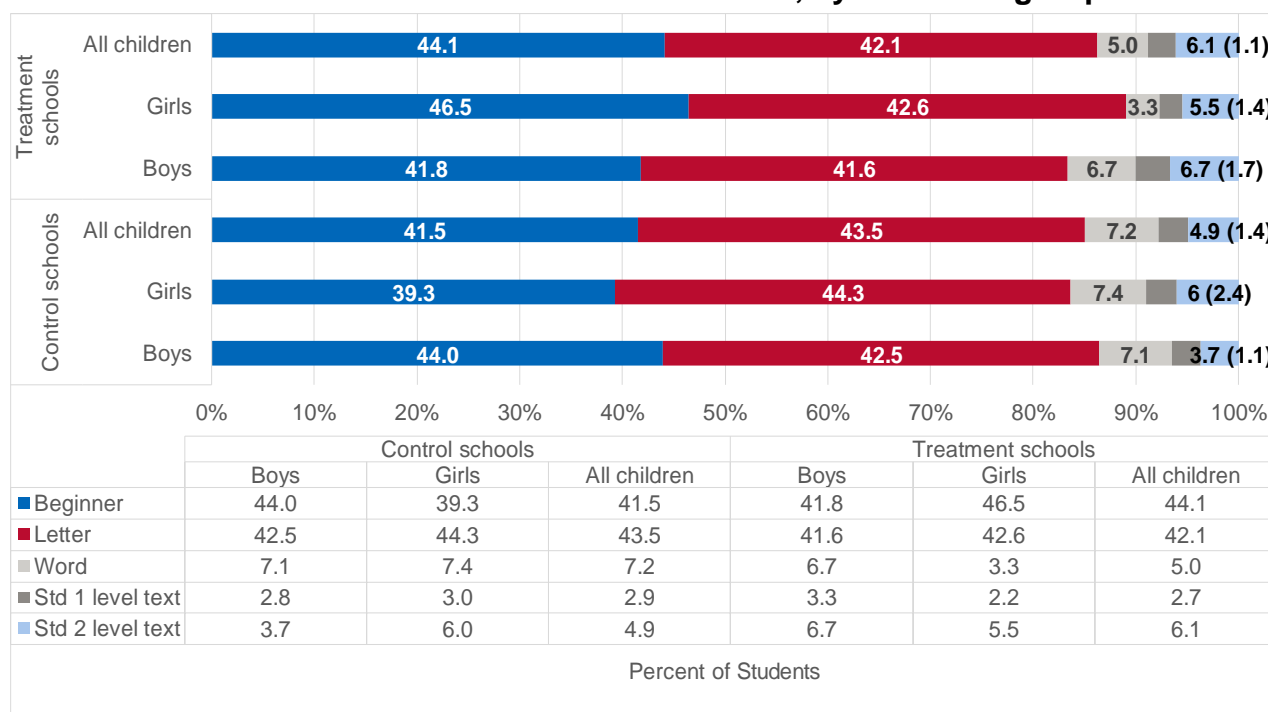
**Table 29.    Student demographics for Start Early: Read in Time—Uttar Pradesh, by treatment group**

|  | Control Mean (Standard Error) | Treatment Mean (Standard Error) |
|---|---|---|
| Breakfast before school | 83.3% (6.3) | 90.2% (1.6) |
| Help with homework at home | 52.8% (4.3) | 54.2% (3) |
| Extra reading material at home | 36.4% (4.8) | 35.4% (2.9) |
| Absent from school in last week | 42.3% (6) | 46.7% (3.1) |
| Someone reads aloud at home | 41.4% (5.6) | 38.4% (2.8) |

### 6.4.2   ASER Results

**Figure 26** below shows performance of Standard 2 students in treatment and control schools for Start Early: Read in Time in Uttar Pradesh. Of students in treatment schools, 3.8% could read Standard 2 level and another 2.6% of children could read Standard 1-level text but not Standard 2-level text. In control schools, the proportion of children who could read Standard 2-level text was almost similar to that in treatment schools (at 3.6%). About 56% of students in treatment schools and approximately 59% of students in control schools could not read letters.

**Figure 26.    Percent of students by ASER reading level for Start Early: Read in Time—Uttar Pradesh, by treatment group**



| | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|
| | Boys | Girls | All children | Boys | Girls | All children |
| ■ Beginner | 59.4 | 58.3 | 58.8 | 54.8 | 57.1 | 55.9 |
| ■ Letter | 32.2 | 30.3 | 31.2 | 34.9 | 34.7 | 34.8 |
| ■ Word | 3.1 | 3.7 | 3.4 | 2.4 | 3.6 | 3.0 |
| ■ Std 1 level text | 2.9 | 3.1 | 3.0 | 3.8 | 1.2 | 2.6 |
| ■ Std 2 level text | 2.4 | 4.6 | 3.6 | 4.1 | 3.4 | 3.8 |

Percent of Students
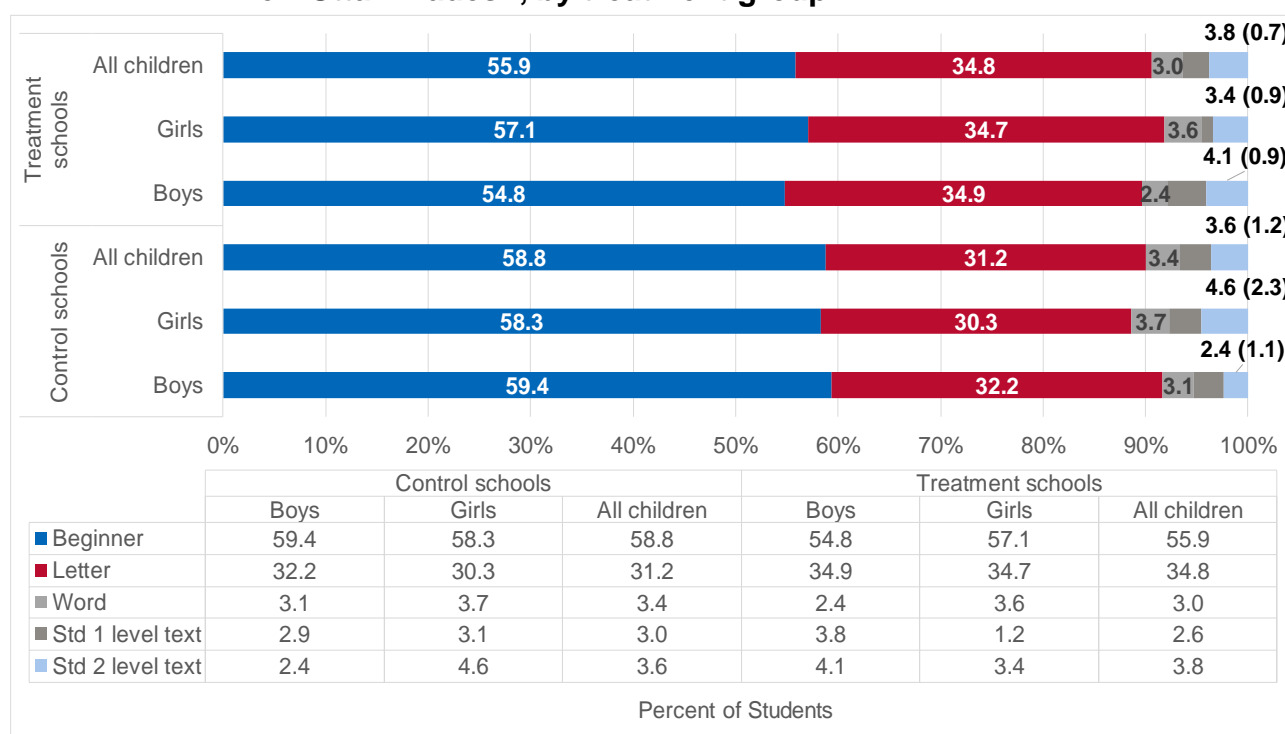
### 6.4.3   EGRA Results

The mean scores and effect sizes in ORF and reading comprehension for Standard 2 students by treatment group in Start Early: Read in Time in Uttar Pradesh are shown in **Table 30**.

**Table 30.    Mean ORF and reading comprehension score by treatment group (Start Early: Read in Time—Uttar Pradesh)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 2.4 (0.6) | 0.00 |
| | Treatment | 2.4 (0.4) | |
| Reading comprehension (% correct) | Control | 2.5% (0.9) | 0.03 |
| | Treatment | 2.8% (0.5) | |

Performance was nearly identical across control and treatment groups, with students in both groups scoring 2.4 cwpm. Poor performance in reading comprehension is to be expected based on students' low ORF scores; roughly 95% of students could not answer a single reading comprehension question. Not surprisingly, the mean scores were less than 3% for both groups. If we consider the groups to be "balanced," then these results suggest that there has been no impact from the intervention in Standard 1 on student performance on oral reading at the beginning of Standard 2.

These low scores might be explained by the population of the sample, who were children from marginalized communities. As referenced in **Annex B**, these students were typically

first-generation learners, which may mean there is a lack of support for learning from the community.

*Figure 27* below shows the distribution of ORF by control and treatment groups. The distribution for both groups is heavily skewed to the right, with majority of students unable to recognize a single word and, therefore, scoring zero. This graph reinforces the lack of difference between the two groups.

**Figure 27.   Distribution of ORF, by treatment group (Start Early: Read in Time—Uttar Pradesh)**
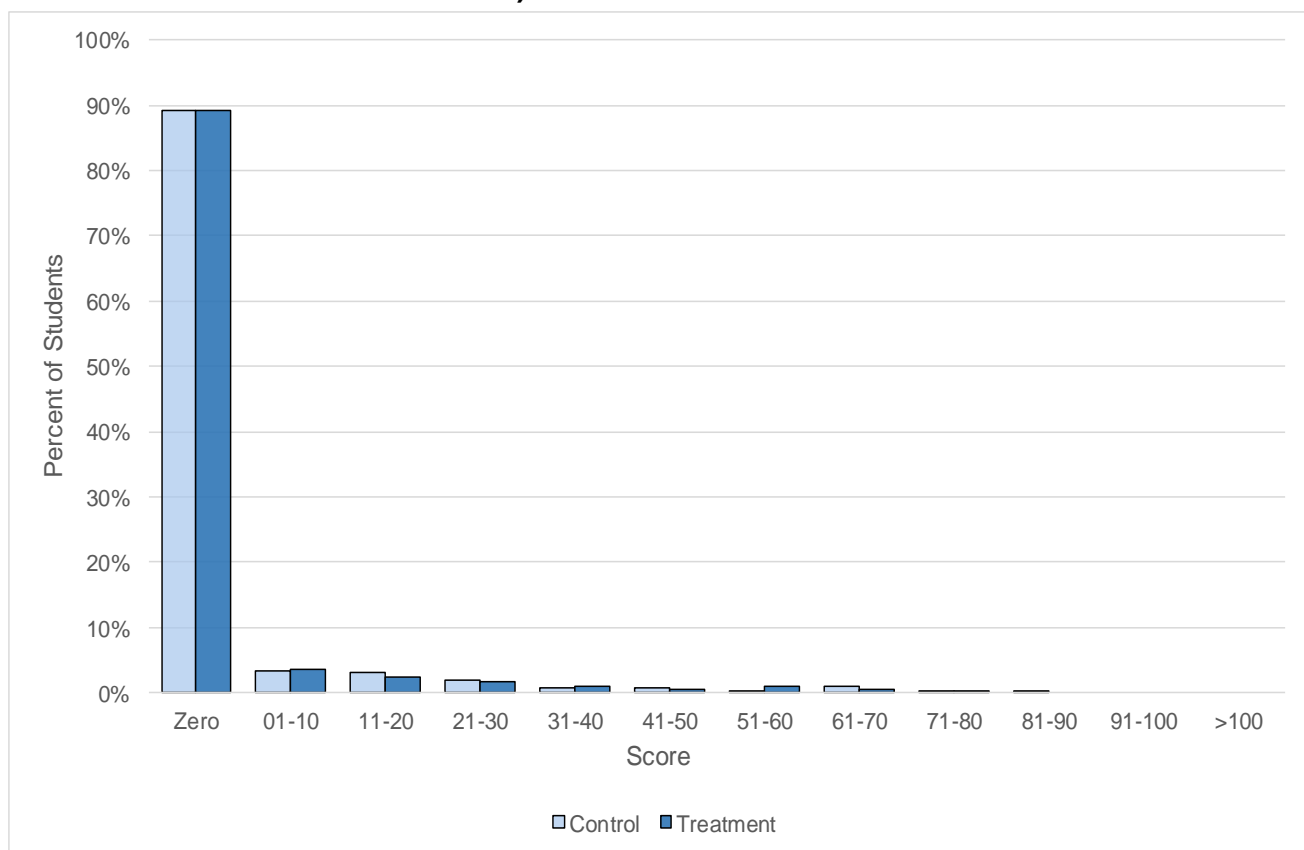


### 6.4.4   ORF by ASER Categories

*Table 31* presents the range of ORF scores per each of the ASER reading categories. The majority of students in this project location were categorized as beginner or letter reading level, with a median ORF of 0 cwpm. The small number of students reading Standard 2-level text were reading about 40 cwpm.

**Table 31.    Range of ORF per ASER reading category (Start Early: Read in Time—Uttar Pradesh)**

| ASER Reading Category | Oral Reading Fluency | | | | | | |
|---|---|---|---|---|---|---|---|
| | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0.03 (0.37) | 59% | 0 | 0 | 0 | 0 | 7 |
| Letter | 0.46 (2.53) | 32% | 0 | 0 | 0 | 0 | 39 |
| Word | 5.85 (6.36) | 3% | 0 | 0 | 3 | 8 | 28 |
| Standard1 | 17.96 (9.99) | 3% | 0 | 11 | 14 | 28 | 62 |
| Standard 2 | 41.57 (21.13) | 4% | 0 | 23 | 40 | 65 | 88 |

### 6.4.5   Start Early: Read in Time – Odisha

For the Start Early: Read in Time project in Odisha, 53% percent of the students assessed in the control group were girls and 47% were boys. In treatment schools, an equal number of boys and girls were assessed. The majority of students in both 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 at home, which was the language of the assessment.

*Figure 28* presents the distribution of students' SES in control and treatment schools. Nearly 37% of students in both treatment and control schools in Odisha were categorized as mid SES; however, there were significantly more students with high SES in treatment schools than in control schools (43% and 26%, respectively), and control schools had more students with low SES compared to treatment schools (37% and 20%, respectively).

**Figure 28.    Student SES distribution for Start Early: Read in Time—Odisha, by treatment group**
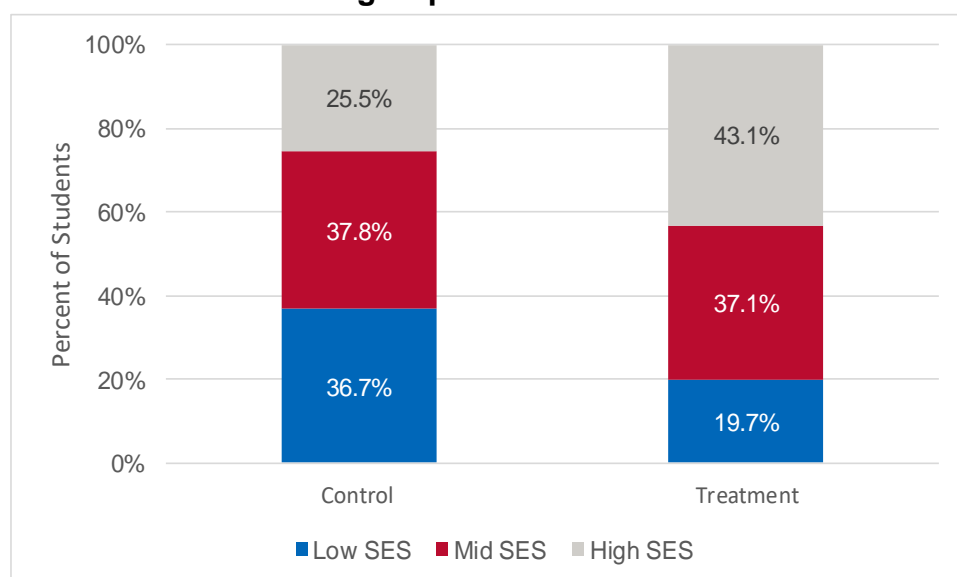


As shown in *Table 32*, in the treatment group, more students from treatment schools reported having breakfast before school, having help with homework at home, having access

to reading materials at home, and having someone read aloud at home. Additionally, the percent of students absent from school in the previous week was lower for students in the treatment schools than for those in control schools.

**Table 32.    Student demographics for Start Early Read in Time—Odisha, by treatment group**

|  | **Control**<br>**Mean (Standard Error)** | **Treatment**<br>**Mean (Standard Error)** |
|---|---|---|
| Breakfast before school | 92.0% (1.7) | 95.8% (1.2) |
| Help with homework at home | 68.3% (3.7) | 83.5% (2.2) |
| Extra reading material at home | 43.6% (3.8) | 52.7% (3.6) |
| Absent from school in last week | 55.2% (3.3) | 46.5% (3.7) |
| Someone reads aloud at home | 43.0% (3.5) | 55.9% (3.3) |

### 6.4.6   ASER Results

***Figure 29*** shows performance of Standard 2 students in treatment and control schools. Of students in treatment schools, 22.3% could read Standard 2-level text and another 12.9% of students could read Standard 1-level text, but not Standard 2-level text. In control schools, only 14.7% of students could read Standard 2-level text. Regression analysis (see ***Annex I***) confirms that the percent of students who could read Standard 2-level text in treatment schools was significantly higher than in control schools; however, this difference disappears when we control for child and household characteristics.

**Figure 29.    Percent of students by ASER reading level for Start Early: Read in Time—Odisha, by treatment group**



| | | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|---|
| | | Boys | Girls | All children | Boys | Girls | All children |
| ■ | Beginner | 31.0 | 30.4 | 30.7 | 20.4 | 23.8 | 22.1 |
| ■ | Letter | 37.4 | 45.4 | 41.7 | 35.5 | 31.3 | 33.4 |
| ■ | Word | 13.0 | 5.0 | 8.7 | 9.1 | 9.6 | 9.3 |
| ■ | Std 1 level text | 5.5 | 3.1 | 4.2 | 14.5 | 11.3 | 12.9 |
| ■ | Std 2 level text | 13.2 | 16.1 | 14.7 | 20.5 | 24.1 | 22.3 |

Percent of Students
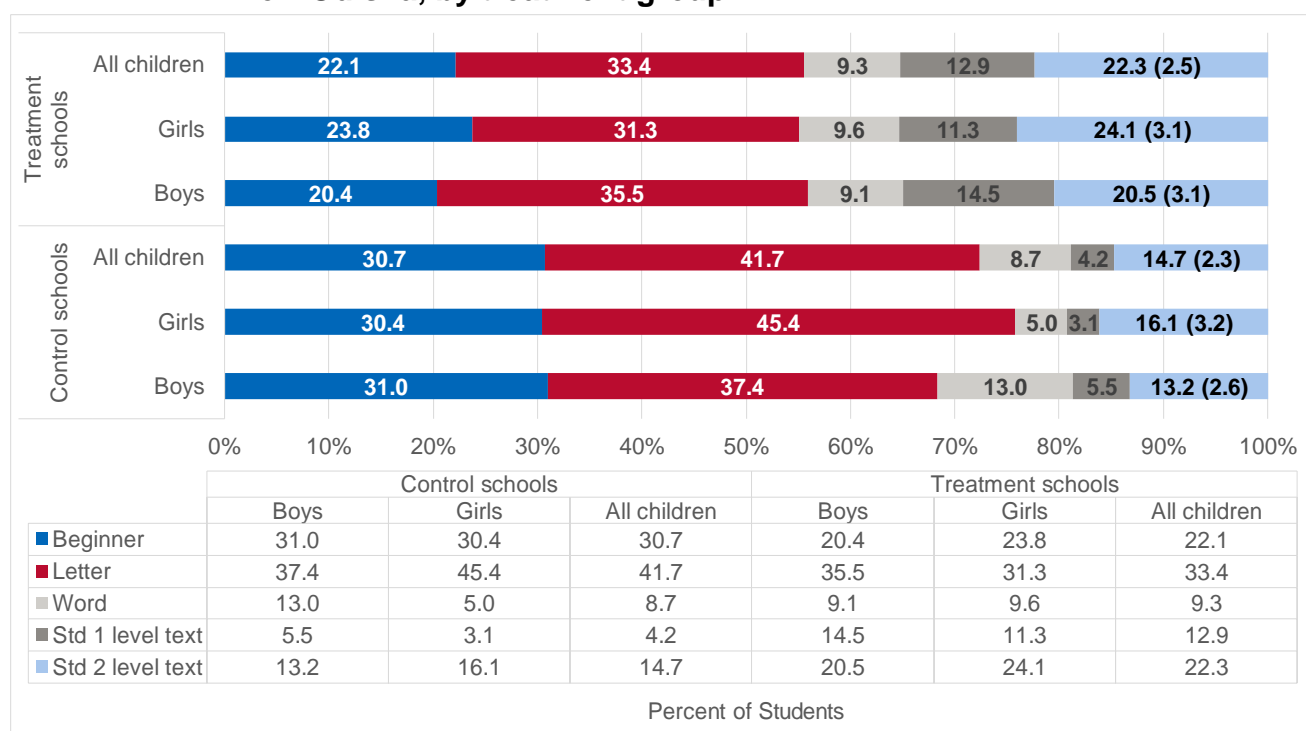
## *6.4.7 EGRA Results*

**Table 33** below displays mean scores and effect sizes in ORF and reading comprehension for Standard 2 students by treatment group in Start Early: Read in Time in Odisha.

**Table 33.  Mean ORF and reading comprehension score by treatment group (Start Early: Read in Time—Odisha)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 5.8 (0.8) | 0.11 |
| | Treatment | 10.8 (1.4) | |
| Reading comprehension (% correct) | Control | 6.5% (0.9) | 0.07 |
| | Treatment | 10.3% (1.6) | |

Performance across control and treatment groups was significantly different, with students in control schools reading 5.8 cwpm and students in treatment schools reading 10.8 cwpm. However, the corresponding effect size of 0.11 is not classified as a small effect. Due to students' difficulties in reading the passage, they also scored low in reading comprehension; roughly 82% of students in the control group could not answer a single reading comprehension question, while 73% of students in the treatment group could not answer a single reading comprehension question.

The performance difference between control and treatment groups might be explained in part by the socioeconomic and demographic characteristics of the sample. As discussed earlier, 43.1% of students in the treatment group belong to the highest socioeconomic level, while only 25.5% of students in the control group belong to the higher socioeconomic group. Hence, the balance between control and treatment groups within this sample is questionable.

**Figure 30** below shows the distribution of ORF by control and treatment groups. The distribution for both groups is heavily right-skewed, with a majority of students unable to recognize a single word and scoring zero.

**Figure 30.     Distribution of ORF, by treatment group (Start Early: Read in Time—Odisha)**
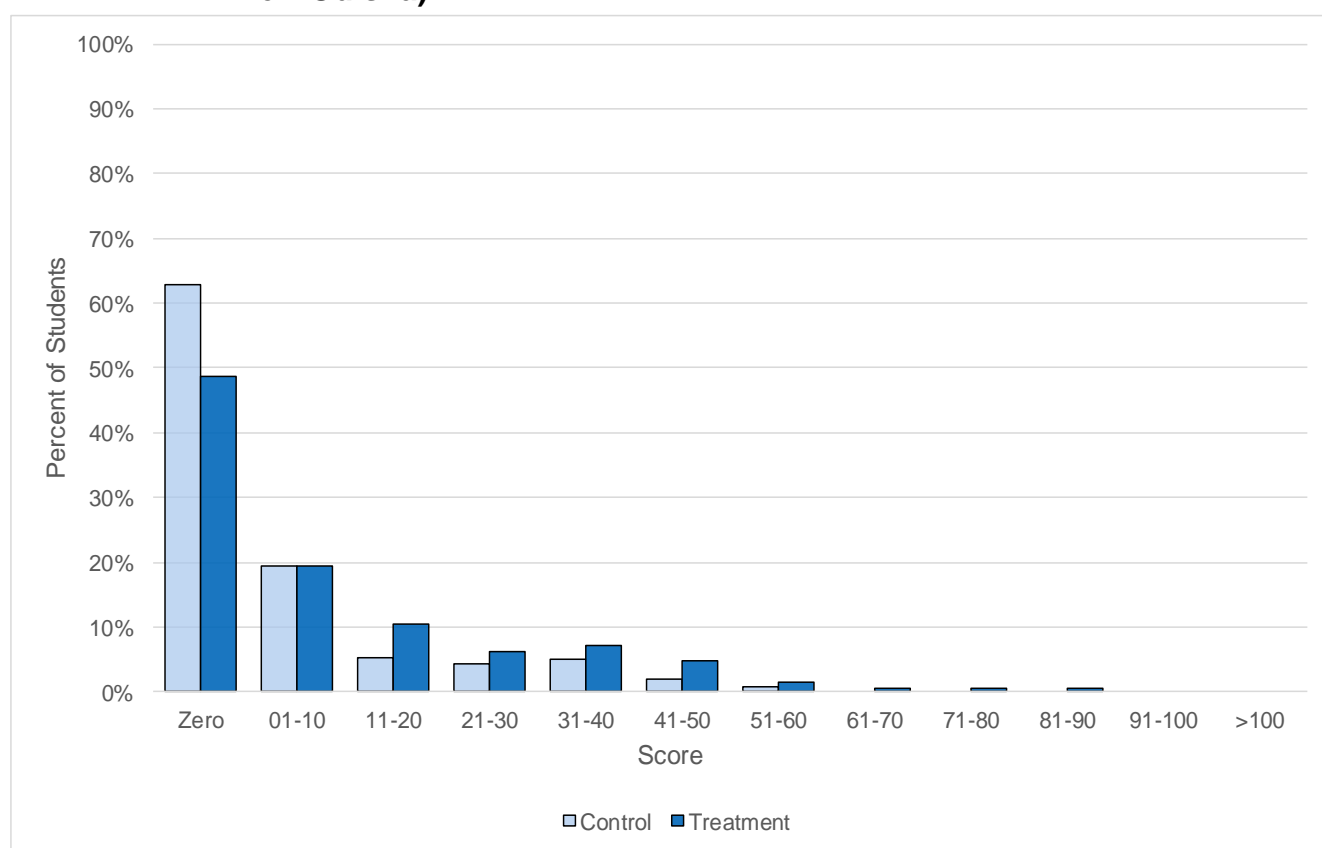


This graph supports the findings in ***Table 33*** that a difference exists between control and treatment groups. There were slightly more students in the control group (approximately 63%) who were unable to read a single word compared to students in the treatment group (approximately 49%). Additionally, more students in the treatment schools performed better in each category range above the 11–20 cwpm category.

### *6.4.8   ORF per ASER Category*

***Table 34*** presents the range of ORF scores per each of the ASER reading categories. The majority of students in this project location were categorized at the letter reading level, reading between 0 and 15 cwpm. Those students who were able to read Standard 2-level text were reading about 33 cwpm.

**Table 34.     Range of ORF per ASER reading category (Start Early: Read in Time—Odisha)**

| ASER Reading Category | Oral Reading Fluency | | | | | | |
|---|---|---|---|---|---|---|---|
| | Estimate (SD) | Percent | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0.03 (0.3) | 25% | 0 | 0 | 0 | 0 | 6 |
| Letter | 0.7 (1.96) | 37% | 0 | 0 | 0 | 0 | 15 |
| Word | 6.95 (4.93) | 9% | 0 | 4 | 7 | 9 | 27 |
| Standard1 | 15.68 (11.55) | 10% | 0 | 8 | 11 | 20 | 55 |
| Standard 2 | 33.56 (14.98) | 19% | 9 | 23 | 33 | 43 | 86 |

## 6.5 Right to Read

English Helper's Right to Read project was implemented from September 2015 to September 2017 in Maharashtra, Gujarat, Delhi, Tamil Nadu, Telangana, Karnataka, and Punjab. The project reaches over one million students in Standards 1 to 8 across eight states in India. The 2-year project was extended by an additional year to 100 schools in Maharashtra and expanded to 300 new schools in West Bengal in 2017. Only schools in 6 districts were purposively sampled out of the 16 districts in which the program is being implemented in the state. These 6 districts have more than 75% of the intervention schools. The initial data were collected in the first year of project implementation. Learners were assessed in English.

### 6.5.1 Right to Read – Maharashtra

For the Right to Read project in Maharashtra, 49% of the students assessed in control schools were girls and 51% were boys. In the treatment group, 51% of students were girls and 49% were boys. Most students from control and treatment schools indicated that they speak Marathi at home (about 65%). Additionally, a large percent of students indicated that they speak Hindi at home (17% of students in control schools and 22% in treatment schools).

**Figure 31** shows just over one-third of students were categorized as mid-SES in both treatment and control schools. Treatment schools had slightly more students with high SES (22% in control and 36% in treatment), whereas control schools had more students categorized as having low SES.

**Figure 31.    Student SES distribution for Right to Read—Maharashtra, by treatment group**
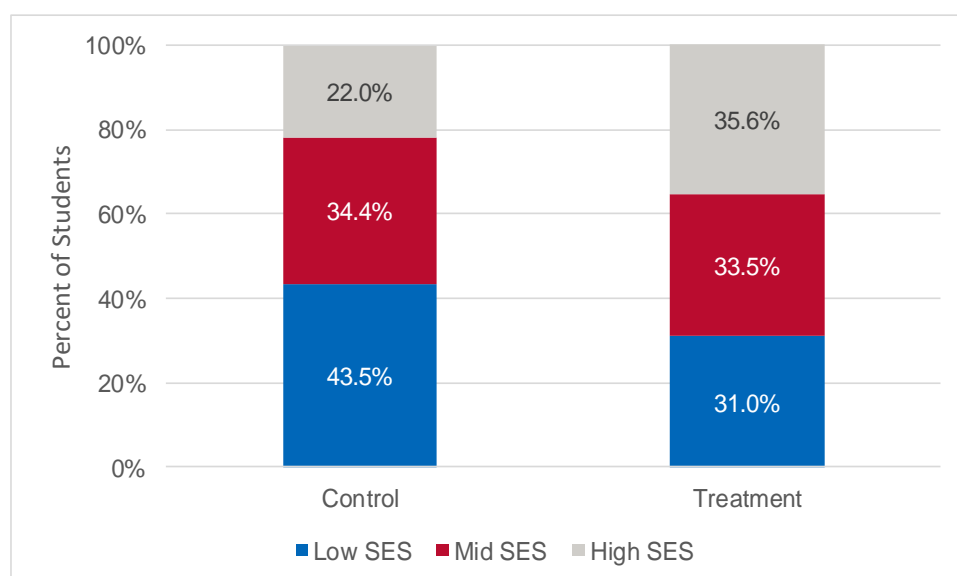


**Table 35** shows that more students in treatment schools indicated having breakfast at home, having help with homework at home, having extra reading material at home, and having someone read aloud at home. Relatively few students from this project location reported being absent from school in the previous week (about 36%) when compared to students from other project locations.

**Table 35.    Student SES distribution for Right to Read—Maharashtra, by treatment group**

| | Control<br>Mean (Standard Error) | Treatment<br>Mean (Standard Error) |
|---|---|---|
| Breakfast before school | 78.4% (2.4) | 83.5% (1.5) |
| Help with homework at home | 67.4% (3.8) | 75.2% (2) |
| Extra reading material at home | 41.9% (4.3) | 44.1% (2.1) |
| Absent from school in last week | 37.2% (6.1 | 36.4% (2.1) |
| Someone reads aloud at home | 58.1% (3.2 | 67.5% (1.9) |

### 6.5.2    ASER Results

Right to Read's project in Maharashtra works to improve the English ability of students, and therefore, students were tested using an English reading tool. ***Figure 32*** below shows Standard 2 students' performance in treatment and control schools. Fewer than 10% of students across both treatment and control schools could read words or higher on the ASER reading tool. High percentages of students, 42.6% in treatment schools and 44.9% in control schools, were unable to read letters.

**Figure 32.    Percent of students by ASER reading level for Right to Read—Maharashtra, by treatment group**



| | Control schools | | | Treatment schools | | |
|---|---|---|---|---|---|---|
| | Boys | Girls | All children | Boys | Girls | All children |
| ■ Beginner | 42.9 | 47.0 | 44.9 | 43.9 | 41.4 | 42.6 |
| ■ Letter | 50.6 | 47.8 | 49.2 | 50.2 | 52.8 | 51.5 |
| ▪ Word | 6.1 | 4.5 | 5.4 | 5.0 | 5.0 | 5.0 |
| ▪ Std 1 level text | 0.4 | 0.7 | 0.6 | 1.0 | 0.6 | 0.8 |
| ▪ Std 2 level text | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.1 |

Percent of Students
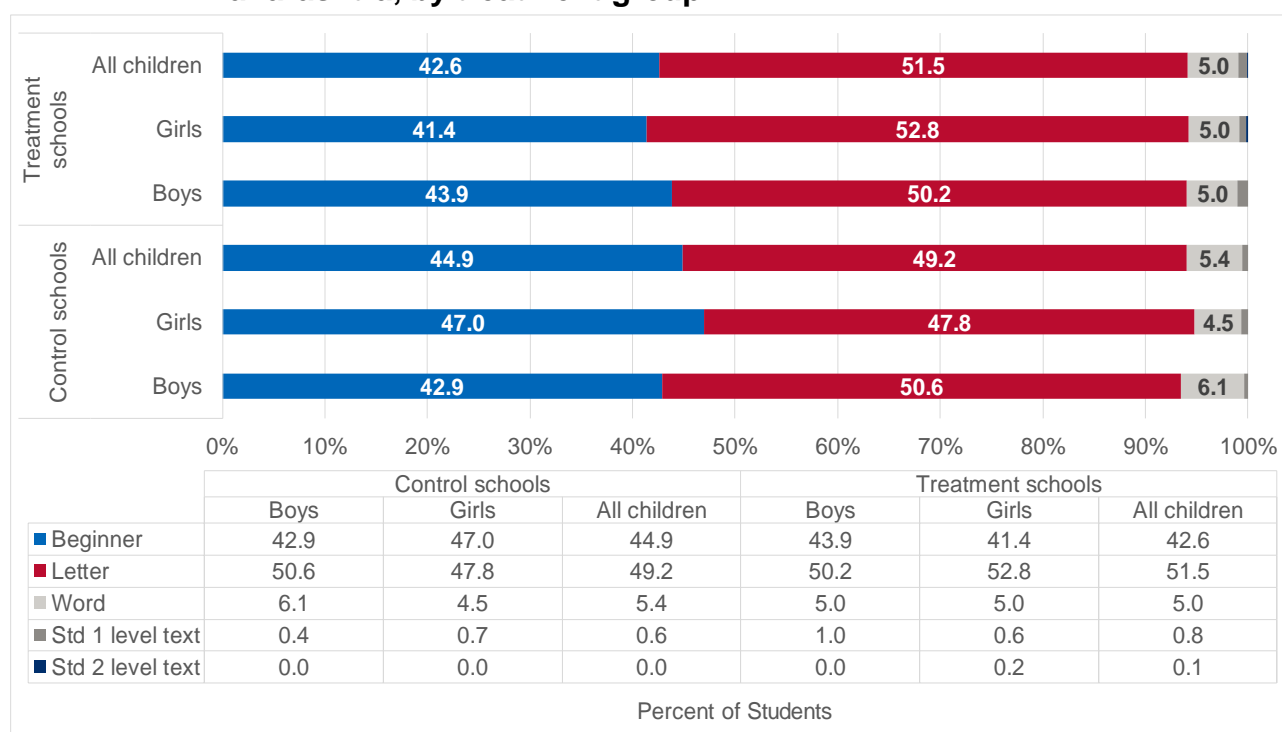
### 6.5.3    EGRA Results

***Table 36*** below displays mean scores and effect sizes in ORF and reading comprehension for Standard 2 students by treatment group in the Right to Read project in Maharashtra.

**Table 36.     Mean ORF and reading comprehension score by treatment group (Right to Read—Maharashtra)**

| Mean Scores | Treatment Group | Mean & Standard Error | Effect Size |
|---|---|---|---|
| Oral reading fluency (ORF) | Control | 0.6 (0.2) | 0.09 |
| | Treatment | 1.0 (0.2) | |
| Reading comprehension (% correct) | Control | 0.1% (0.1) | 0.11 |
| | Treatment | 0.6% (0.2) | |

Performance was similar across control and treatment groups; students in control schools read 0.6 cwpm, and students in treatment schools read 1.0 cwpm[9]. Given that students were tested in English, it is clear that students were struggling in a language that was not their first language; most students could not recognize a single word of English.

*Figure 33* below shows the distribution of ORF by control and treatment groups. The distribution for both groups is heavily right skewed, with a majority of students (over 90%) unable to recognize a single word, and therefore, scoring zero.

**Figure 33.     Distribution of ORF, by treatment group (Right to Read—Maharashtra)**



---
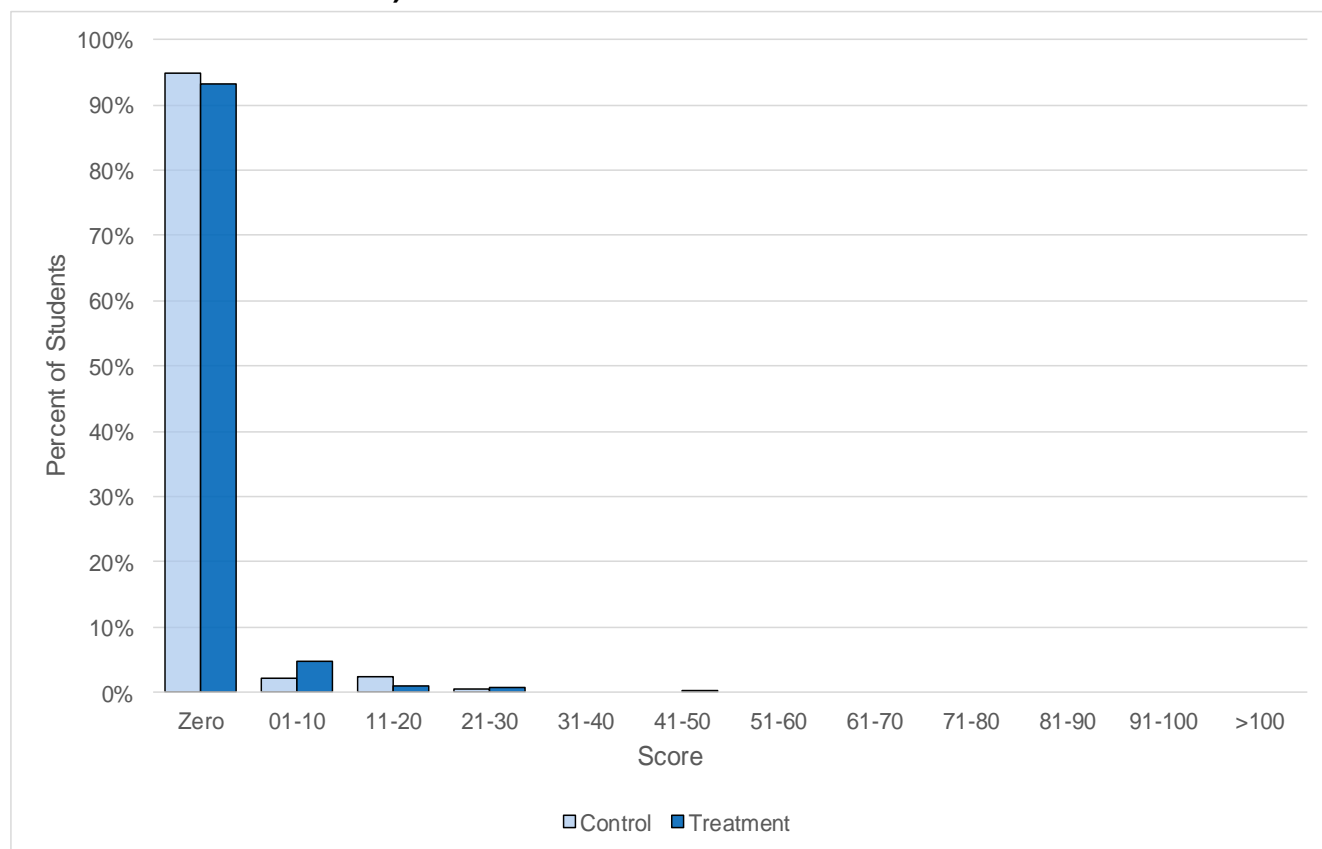
[9] Note: One treatment school's data was removed from analysis. The ORF mean for the school was found to be many standard deviations outside the overall mean for the treatment group. Upon further investigation, the school was an English medium school which explained the unusually high performance.

### 6.5.4 ORF per ASER Category

**Table 37** presents the range of ORF in terms of cwpm for each of the ASER reading categories. The majority of students in this project location were categorized at the beginner and letter levels, with median fluency of 0 cwpm. Those students who were able to read connected text and were categorized at the Standard 1 reading level were only reading about 17 cwpm.

**Table 37.    Range of ORF per ASER reading category (Right to Read—Maharashtra)**

| ASER Reading Category | Oral Reading Fluency | | | | | |
|---|---|---|---|---|---|---|
| | Estimate (SD) | Minimum | Quartile 2 | Median | Quartile 3 | Maximum |
| Beginner | 0 (0) | 0 | 0 | 0 | 0 | 0 |
| Letter | 0.21 (1.43) | 0 | 0 | 0 | 0 | 13 |
| Word | 9.63 (7.31) | 0 | 5 | 10 | 14 | 24 |
| Standard 1 | 15.01 (8.47) | 0 | 17 | 17 | 17 | 44 |
| Standard 2* | 46 (0) | * | * | * | * | * |

* Data are not shown here due to minimal observations.

# 7 Early Grade Reading Benchmarks

## 7.1 Introduction to Benchmarking

The primary purpose of the benchmarking activity was to set reading proficiency benchmarks in five languages used in USAID/India reading projects. The benchmarks will be used to count the number of students in each project who read at the benchmark level or above. This will enable the results of the reading projects to be included in the Global Count. It will also enable comparisons of reading results across projects. The data will additionally be used to report progress on standard USAID Indicator E.S.1-1: "The percent of learners who demonstrate reading fluency and comprehension of grade-level text at the end of Grade 2 with USG assistance." However, this latter requirement is not possible without first defining standards/benchmarks that represent "reading fluency and comprehension" for each of the five languages represented in these data: Hindi, Kannada, Marathi, Oriya, and English.

Accordingly, a benchmarking workshop was held in New Delhi during the first week of May 2018. The goal of this workshop was to set reading benchmarks for each language (i.e., a standard measure of whether a learner had acquired proficiency in reading). The emphasis was not just on the mechanics of reading, but also on developing the ability to read with comprehension. Research indicates that in order to read with comprehension, children must possess sufficient skill in both the mechanics of reading (e.g., concepts of print, associating letters and sounds, and ultimately decoding) and understanding or meaning making (e.g., vocabulary, syntax, background knowledge, and ultimately language comprehension). These skills are captured in assessments of students' ORF, which measures the ability to read with accuracy, speed, and expression. ORF is often referred to as the bridge that connects word decoding and comprehension. As such, it has been found to be highly predictive of reading comprehension. Since comprehensive, reliable measures of reading comprehension are difficult to develop (particularly for emerging readers), ORF benchmarks have been used across many USAID programs to provide estimates for "reading fluency and comprehension" based on the close relationship between ORF and reading comprehension.

The process of benchmarking is highly dependent on the country context, languages in consideration, and the government system (curriculum, reading strategy, etc.). Therefore, while the benchmarking workshop was led by RTI benchmarking experts, final benchmarks were ultimately determined by participants with direct knowledge of languages and context in question, such as in-country project staff, linguists, curriculum specialists, government representatives, and reading specialists. The workshop was attended by 43 participants, representing 20 organizations across India and internationally.

As an introduction to the benchmarking process, RTI experts led discussions on the development of reading, the purposes of assessment, and benchmarking principles. There was also extensive discussion of the similarities and differences in the linguistic characteristics of the five languages used in the projects.

## 7.2 Overview of Benchmarking Process

A simplified overview of the benchmarking process can be seen in ***Figure 34***, below. From the outset, the purpose of setting benchmarks was clear: to provide USAID with a standard against which project performance can be compared, with the express requirement of reporting against E.S. 1-1. Therefore, the benchmarking process in India began with a review of data and analyses, to help participants understand what types of measures were available for benchmarking. This was followed by small group discussions on the appropriate metric to use for benchmarking, for example whether to use a direct measure or a proxy/indirect measure of comprehension. The discussion also addressed whether language-specific or universal benchmarks should be set. These discussions led to group-level decisions being made for metrics and performance thresholds for the benchmark indicator in each language. The final step in the process focused on building consensus

around the benchmarks, in order to ensure that there was clarity in the rationale and consistency in the final measures selected for each project.

**Figure 34.    India benchmarking process**
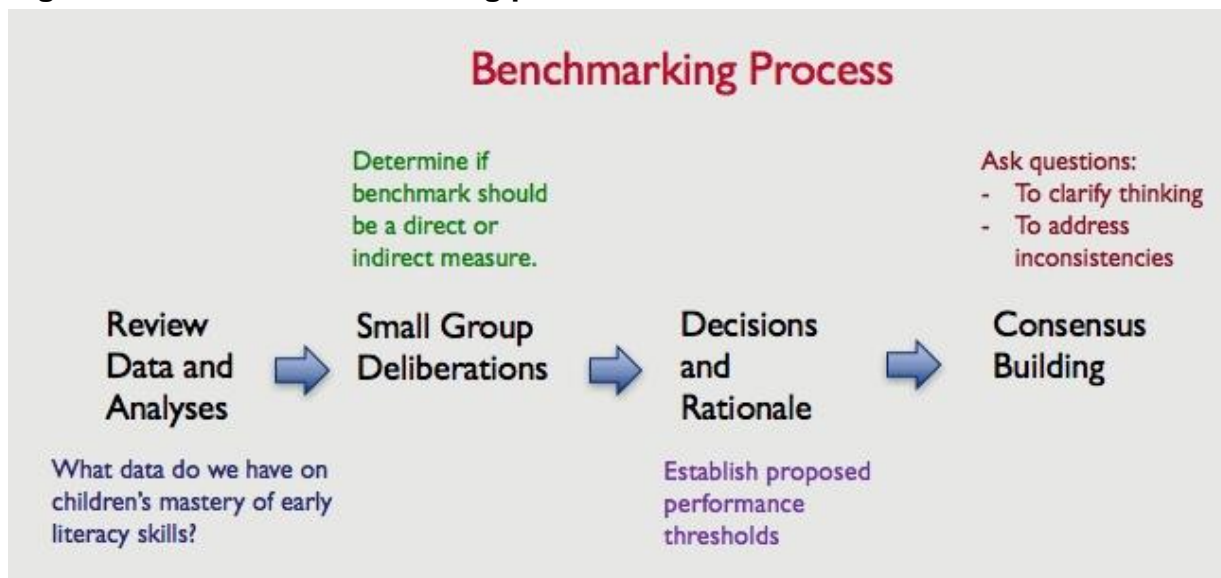


## 7.3    Process in Detail

All analyses presented in the workshop were derived from project baseline data (by language), as measured by the common external assessment. Since there was no a priori assumption about which variable would be used for benchmarking, participants were provided with analyses across all administered subtasks (i.e., ORF, reading comprehension, and ASER performance levels) for review. All cross-subtask relationships were estimated using two separate approaches: 1) mean method (plotting the mean of one subtask against a proficient score on a second subtask); and 2) logistic regression method (estimating the point on the first subtask at which a student has a .5 probability of scoring proficiently on the second subtask).

Upon review of the data, all groups came to the independent conclusion that the most appropriate benchmark for these purposes would be a measure of ORF that constitutes a proficient level of reading comprehension. As such, they selected an indirect measure of reading comprehension (due to the aforementioned difficulties in measuring comprehension reliably). Concern was expressed that oral reading fluency was not a strong proxy for comprehension, particularly in transparent Indian languages. However, participants recognized the benefits of using a fluency measure and agreed that it could justifiably be benchmarked against comprehension. This was seen as preferable to setting a benchmark based on achievement of an ASER level, since ASER does not explicitly include comprehension. Furthermore, when asked to determine whether they felt that a

> **Example of Methods for Estimating Oral Reading Fluency Benchmarks based on Reading Comprehension**
>
> **Mean method**
> Benchmark = mean (average) fluency rate of those with > 80% comprehension
>
> **Logistic Regression method**
> Benchmark = fluency rate where students have a 50% chance of reading with > 80% comprehension

single benchmark should be set for all languages or if each language required an independent benchmark estimate, participants felt that there were sufficient orthographic and

syntactical differences across languages to justify the need for setting a separate benchmark for each language. Participants did not have strong preferences for selecting a benchmarking method (i.e., mean or logistic regression), but instead felt that it was best to use all available information in order to inform their decision about where the ORF benchmark should be set.

Although participants chose to calculate benchmarks for each language separately, the relationships across subtasks and the fluency rates associated with proficiency for both reading comprehension and ASER levels were relatively consistent across languages—as shown in *Figures 35* and *36*.

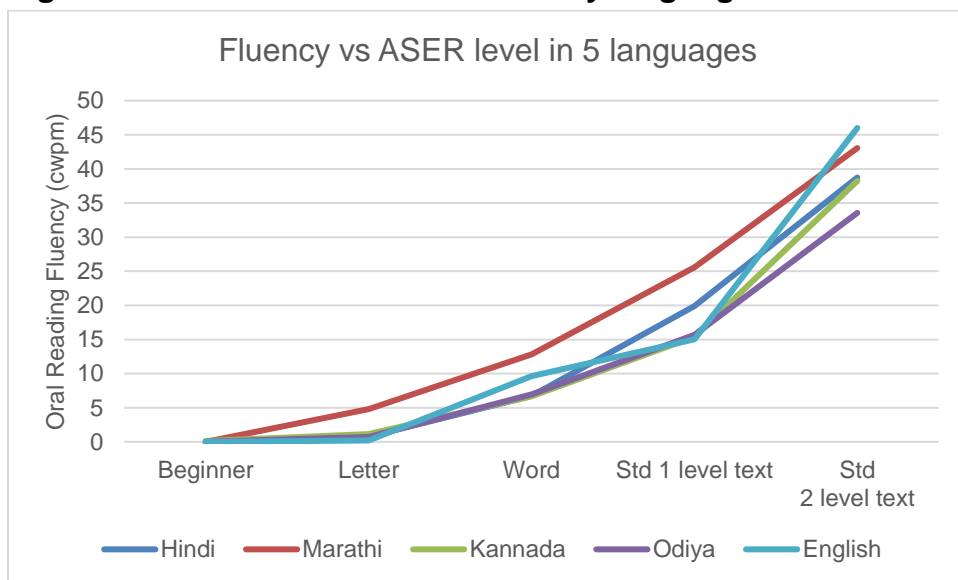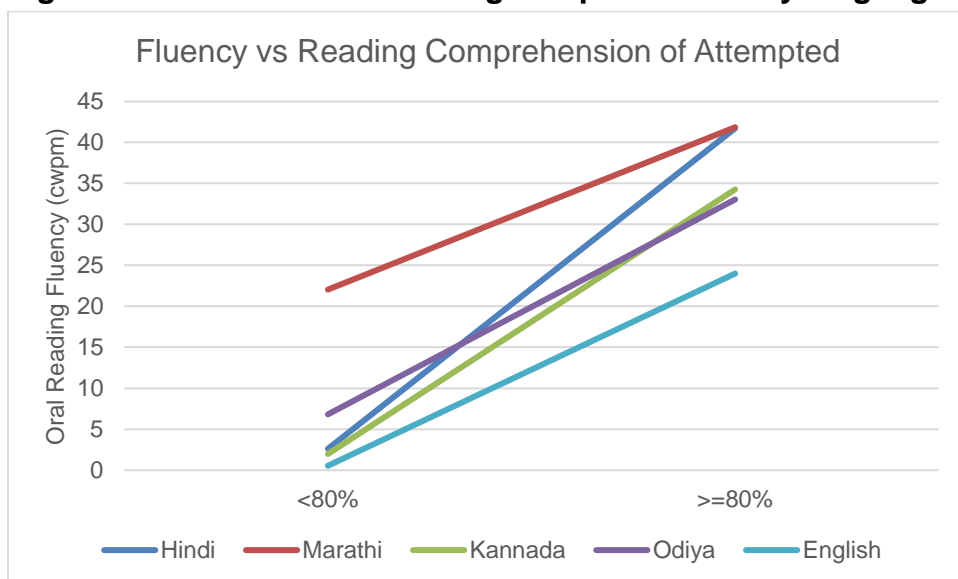**Figure 35.    ORF versus ASER level by language**



**Figure 36.    ORF versus reading comprehension by language**



The following section provides an overview of the data that were presented for each language during the workshop, as well as an explanation for each benchmark that was set.

## 7.4 Results by Language

In this section, results are presented for each of the five languages assessed. A single table and graph are provided for each language. Each table provides the ORF estimate that is associated with each of four outcomes for both methods (logistic regression and mean). The four outcomes are as follows:

1. ASER Standard 1 Level: students who performed at the sentence reading level task or above on ASER
2. ASER Standard 2 Level: students who performed at the paragraph reading level on ASER
3. Reading Comprehension 80% Overall: students who were able to correctly answer at least 80% of the comprehension questions on the EGRA comprehension subtask (i.e., 4 out of 4; 4 out of 5; 5 out of 5)
4. Reading Comprehension 80% of Attempted: students who were able to correctly answer at least 80% of the comprehension questions that they attempted on the EGRA comprehension task (i.e., 2 out of 2; 3 out of 3; 3 out of 4; 4 out of 4; 4 out of 5; 5 out of 5).[10] This measure accounts for the fact that slower readers who do not finish the passage can still meet the required proficiency level of comprehension.

Additionally, the lower and upper bound estimates (calculating using a 95% confidence interval) are presented for each estimate in all tables. For the logistic regression approach, an $R^2$ value is also provided, as estimate of the strength of the relationship (i.e., the percent of variance in the outcome explained by oral reading fluency). The final column of each table lists the number of students who reached the standard (or outcome variable). While the sample size was sufficiently large to calculate estimates for the majority of outcomes across the five languages, in certain instances reliable estimates could not be produced—indicated by "n/a." This was particularly problematic for English, for which only 1 student reached ASER Level 2 and/or Reading Comprehension 80% of Attempted.

The figure that is presented for each language provides similar information to that of the tables, although the outcome of Reading Comprehension 80% Overall has been removed. This decision was made both because of the imprecise estimates available (due to small sample sizes across nearly all languages), as well as the fact that Reading Comprehension 80% of Attempted was deemed more appropriate since it did not artificially penalize students who were unable to complete the entire passage. The dot in each column of the figures represents the ORF estimate for a particular outcome and method, while the lines extending out from those dots represent the 95% confidence interval and the upper/lower bound estimates. The dashed line on each figure represents the final benchmark set by the workshop group.

These tables and figures represent the primary data used by each language group to set their ORF benchmarks.

### 7.4.1 Hindi

Hindi represented the largest group in the workshop, with four of the five projects (all but English Helper) testing students in Hindi. *Table 38* shows that ORF estimates for Hindi ranged from as low as 15 for ASER Standard 1 Level to as high as 64 for Reading Comprehension 80% Overall. Furthermore, there were reasonably large sample sizes for all estimates, leading to small confidence intervals.

---

[10] Due to rounding, students correctly identifying 3 out of 4 questions on the EGRA comprehension subtask were included in the estimate as reading with 80% comprehension.

**Table 38.** **Hindi ORF estimates**

| HINDI | | | | | | |
|---|---|---|---|---|---|---|
| Outcome | Method | Oral Reading Fluency Level | Lower Bound Estimate (95% confidence) | Upper Bound Estimate (95% confidence) | R² | Number of Students Reaching Standard |
| ASER Standard 1 Level | Logistic | 15 | 13 | 16 | 0.53 | |
| | Mean | 20 | 18 | 21 | | 1462 |
| ASER Standard 2 Level | Logistic | 28 | 26 | 31 | 0.406 | |
| | Mean | 39 | 35 | 43 | | 659 |
| Reading Comprehension 80% Overall | Logistic | 64 | 56 | 76 | 0.396 | |
| | Mean | 51 | 42 | 60 | | 138 |
| Reading Comprehension 80% of Attempted | Logistic | 48 | 42 | 56 | 0.328 | |
| | Mean | 42 | 36 | 47 | | 347 |

*Figure 37* represents the main data points that were used in the benchmark setting process for Hindi. For each outcome, estimates showed similar precision regardless of the model employed (logistic regression or mean). Although it was determined that Reading Comprehension 80% of Attempted was the appropriate measure against which a reading fluency benchmark should be set, the Hindi group also took into consideration the lower level of the Standard 2 ASER estimates. They ultimately opted for a benchmark that fell slightly below the lower bound estimate of the mean method for 80% Reading Comprehension—at 35 cwpm.

**Figure 37.    Data for Hindi benchmarking**
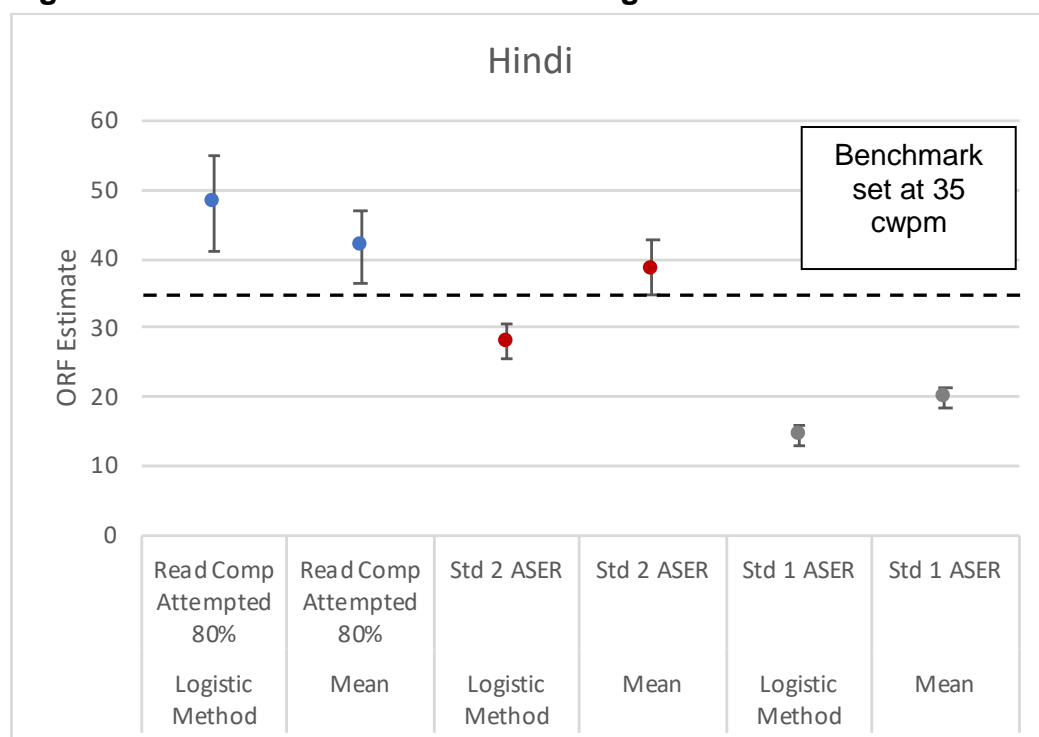


### 7.4.2  *Marathi*

Only the Centre for microFinance (CmF) project administered this assessment in Marathi. Therefore, the sample sizes were smaller than those for Hindi. However, reasonably large sample sizes were still available for all outcomes, resulting in tight confidence intervals. *Table 39* shows that ORF estimates for Marathi ranged from as low as 12 for ASER Standard 1 Level to as high as 59 for Reading Comprehension 80% Overall.

**Table 39.    Marathi ORF estimates**

| MARATHI | | | | | | |
|---|---|---|---|---|---|---|
| **Outcome** | **Method** | **Oral Reading Fluency Level** | **Lower Bound Estimate (95% confidence)** | **Upper Bound Estimate (95% confidence)** | **R²** | **Number of Students Reaching Standard** |
| **ASER Standard 1 Level** | Logistic | 12 | 10 | 13 | 0.458 | |
| | Mean | 26 | 24 | 27 | | 1178 |
| **ASER Standard 2 Level** | Logistic | 28 | 26 | 30 | 0.337 | |
| | Mean | 43 | 41 | 45 | | 762 |
| **Reading Comprehension 80% Overall** | Logistic | 51 | 49 | 53 | 0.56 | |
| | Mean | 59 | 57 | 60 | | 245 |
| **Reading Comprehension 80% of Attempted** | Logistic | 28 | 26 | 29 | 0.212 | |
| | Mean | 42 | 40 | 44 | | 782 |

**Figure 38** represents the main data points that were used in the benchmark setting process for Marathi. While the precision for the logistic and mean method estimates was similar (with tight confidence intervals for both), the estimates were consistently higher for the mean approach. The Marathi group ultimately felt that the logistic estimates were too low – in part because the use of a 0.5 probability level can lead to lower estimates in the logistic regression method. They decided to set their benchmark based on the lower bound of the mean method for Reading Comprehension 80% of Attempted. Therefore, the final benchmark selected for Marathi was 40 cwpm.

**Figure 38.    Data for Marathi benchmarking**



### 7.4.3    Kannada

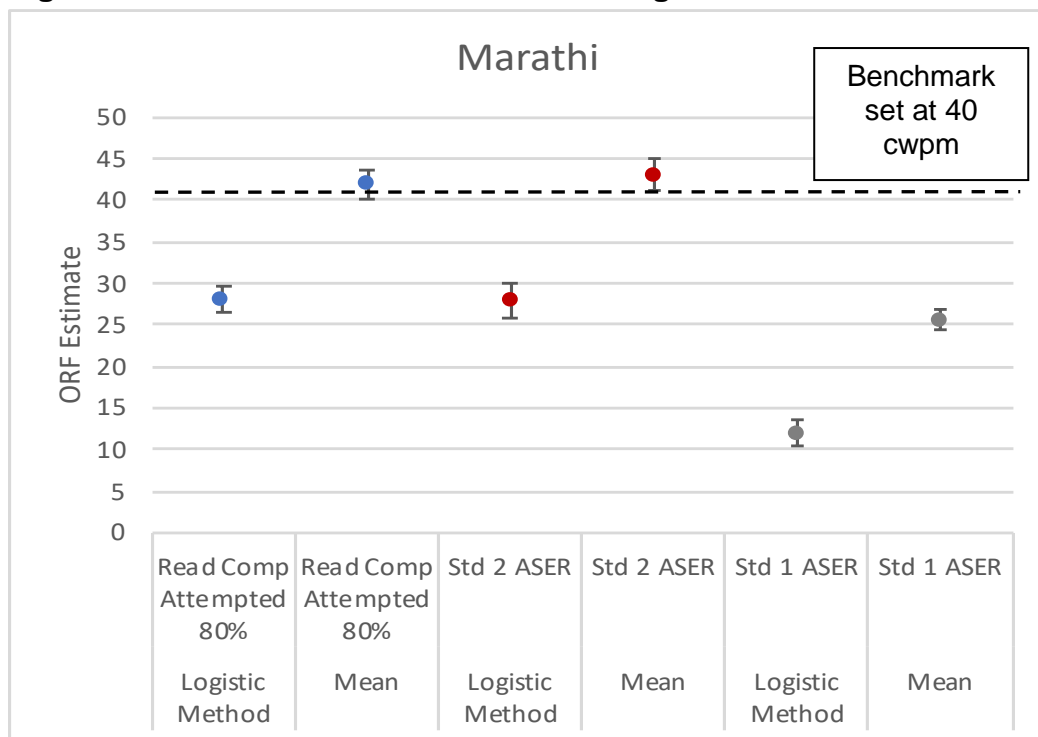Kannada was one of the languages of assessment for the CmF—the only project to assess in this language. **Table 40** shows that ORF estimates for Kannada ranged from as low as 14 for ASER Standard 1 Level to as high as 52 for Reading Comprehension 80% Overall. However, the number of students reaching each outcome was relatively low, which led to large confidence intervals and some unreliable estimates.

**Table 40.     Kannada ORF estimates**

| KANNADA | | | | | | |
|---|---|---|---|---|---|---|
| Outcome | Method | Oral Reading Fluency Level | Lower Bound Estimate (95% confidence) | Upper Bound Estimate (95% confidence) | R2 | Number of Students Reaching Standard |
| ASER Standard 1 Level | Logistic | 14 | 13 | 16 | 0.462 | |
| | Mean | 15 | 13 | 17 | | 113 |
| ASER Standard 2 Level | Logistic | 33 | 28 | 39 | 0.496 | |
| | Mean | 38 | 29 | 47 | | 24 |
| Reading Comprehension 80% Overall | Logistic | n/a | n/a | n/a | n/a | |
| | Mean | 52 | 28 | 76 | | 6 |
| Reading Comprehension 80% of Attempted | Logistic | 39 | 31 | 54 | 0.347 | |
| | Mean | 34 | 24 | 44 | | 17 |

*Figure 39* represents the main data points that were used in the benchmark setting process for Kannada. It is clear from this figure that Kannada had less precise estimates than either Hindi or Marathi. However, the consistency of the estimates across methods was relatively high. Acknowledging that they were working with limited data, the Kannada group ultimately decided to choose a benchmark that fell in the middle of the confidence intervals for both methods using Reading Comprehension 80% of Attempted as the outcome. Therefore, the final benchmark selected for Kannada was 35 cwpm.

**Figure 39.     Data for Kannada benchmarking**
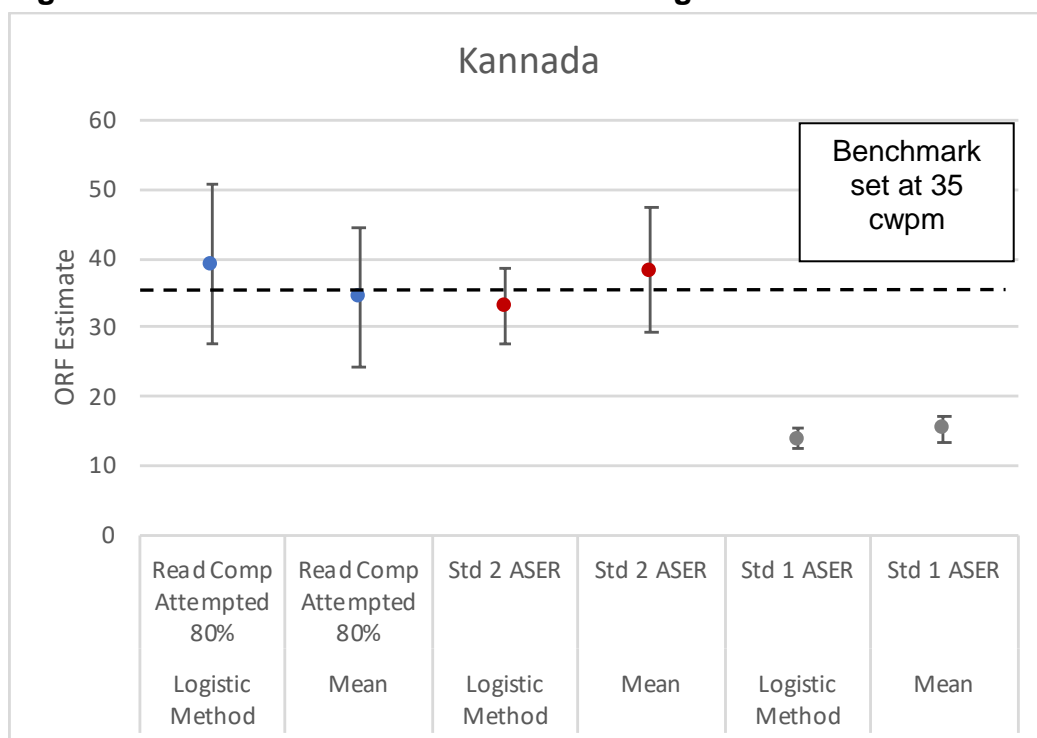
### 7.4.4 Oriya

Oriya was used as a language of assessment only for the CARE India project. ***Table 41*** shows that ORF estimates for Oriya ranged from as low as 10 for ASER Standard 1 Level to as high as 82 for Reading Comprehension 80% Overall. However, the number of students reaching each outcome was relatively low, which led to large confidence intervals, particularly for the EGRA comprehension outcomes.

**Table 41.    Oriya ORF estimates**

| ORIYA | | | | | | |
|---|---|---|---|---|---|---|
| Outcome | Method | Oral Reading Fluency Level | Lower Bound Estimate (95% Confidence) | Upper Bound Estimate (95% Confidence) | R2 | Number of Students Reaching Standard |
| ASER Standard 1 Level | Logistic | 10 | 8 | 11 | 0.578 | |
| | Mean | 16 | 13 | 19 | | 251 |
| ASER Standard 2 Level | Logistic | 21 | 18 | 24 | 0.492 | |
| | Mean | 34 | 31 | 37 | | 166 |
| Reading Comprehension 80% Overall | Logistic | 82 | 64 | 86 | 0.163 | |
| | Mean | | | | | 13 |
| Reading Comprehension 80% of Attempted | Logistic | 52 | 43 | 71 | 0.134 | |
| | Mean | 33 | 28 | 38 | | 66 |

***Figure 40*** represents the main data points that were used in the benchmark setting process for Oriya. It is clear from this figure that Oriya had less precise estimates than either Hindi or Marathi. Additionally, the consistency of the estimates across methods was relatively low. Acknowledging that they were working with limited data, the Oriya group ultimately decided to choose a benchmark that aligned with the mean method using Reading Comprehension 80% of Attempted as the outcome. This estimate was similar to the mean method estimate of Standard 2 ASER. The logistic regression method for 80% Reading Comprehension was higher but was given less weight because of large confidence intervals. Therefore, the final benchmark selected for Kannada was 30 cwpm.

**Figure 40.    Data for Oriya benchmarking**
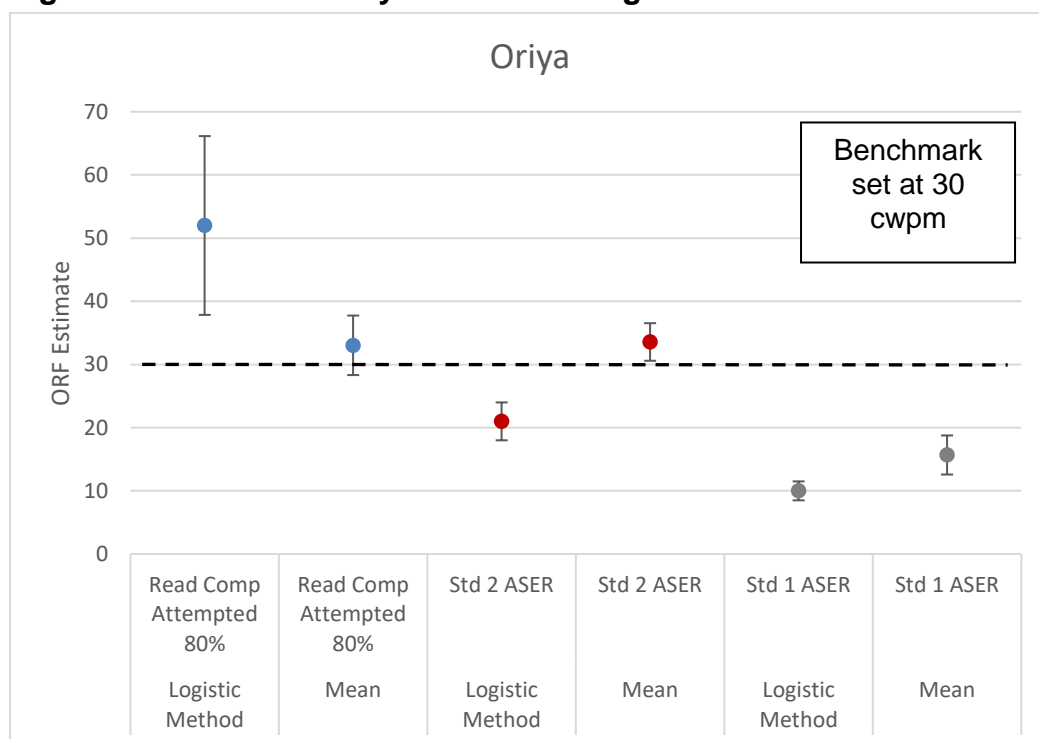


### 7.4.5   English

An English language assessment was administered solely to students from the English Helper project. It is clear from *Table 42* and *Figure 41* that there was insufficient data in the baseline to set an English language benchmark.

**Table 42.    English ORF estimates**

| | | | ENGLISH | | | |
|---|---|---|---|---|---|---|
| **Outcome** | **Method** | **Oral Reading Fluency Level** | **Lower Bound Estimate (95% Confidence)** | **Upper Bound Estimate (95% Confidence)** | **R2** | **Number of Students Reaching Standard** |
| **ASER Standard 1 Level** | Logistic | 30 | 20 | 46 | 0.364 | |
| | Mean | 15 | 10 | 20 | | 18 |
| **ASER Standard 2 Level** | Logistic | n/a | n/a | n/a | n/a | |
| | Mean | 46 | 46 | 46 | | 1 |
| **Reading Comprehension 80% Overall** | Logistic | n/a | n/a | n/a | n/a | |
| | Mean | n/a | n/a | n/a | | 0 |
| **Reading Comprehension 80% of Attempted** | Logistic | n/a | n/a | n/a | n/a | |
| | Mean | 24 | 24 | 24 | | 1 |

**Figure 41.    Data for English benchmarking**
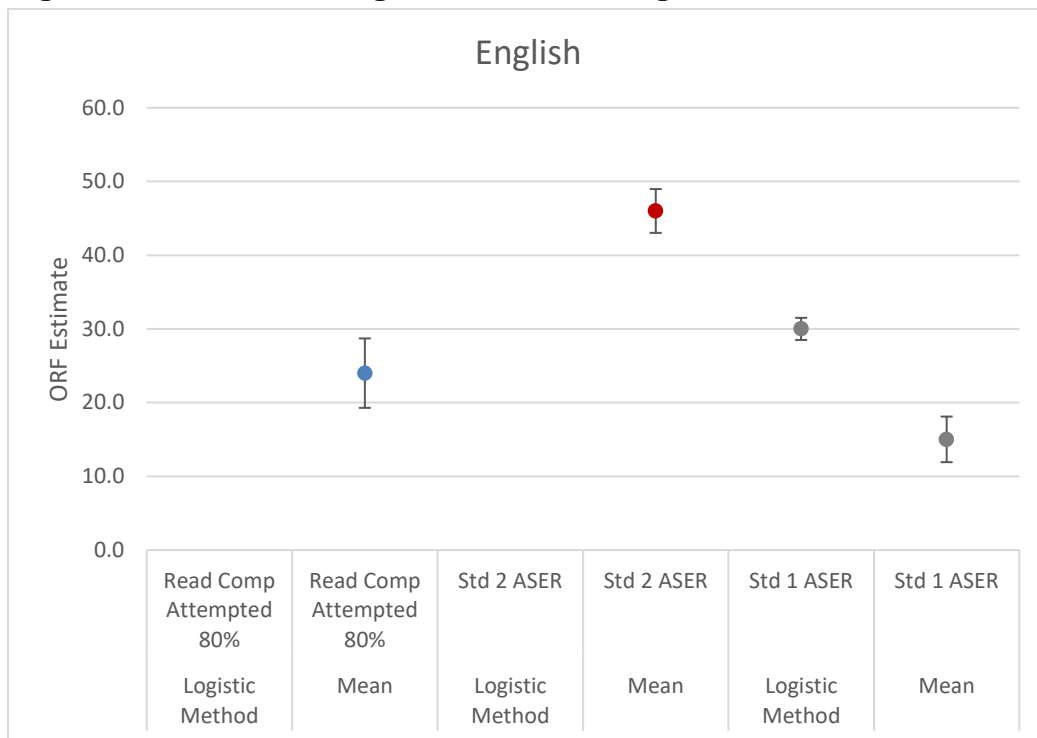


Participants from the English Helper project were provided with data from a variety of other countries in which English benchmarks have been set (in a setting where English is used as a second language). These data consistently showed English language benchmarks of approximately 35 to 45 cwpm.

Endline data were provided to the English Helper project leadership after the conclusion of the workshop, in order to provide further support for setting a benchmark. At endline, the number of students reaching the aforementioned standards was still too low to develop strongly reliable estimates (as only 13 students met the standard of Reading Comprehension 80% of Attempted). Illustratively, the mean ORF scores for the three standards used in each language were as follows: 1) 39 cwpm for ASER Standard 1; 2) 56 cwpm for ASER Standard 2; 3) 31 cwpm for Reading Comprehension 80% of Attempted.

Based on all of these data, the final recommended benchmark for English was set to 30 cwpm.

## 7.5    Final Adopted Benchmarks

While it was ultimately possible to set benchmarks in the four Indian languages (Hindi, Marathi, Kannada, and Oriya) using baseline data, the English benchmark relied on an approach of incorporating information from a range of other countries (and an examination of endline data). Ultimately, the following benchmarks were selected for each of the five languages.

**Table 43.    Final Benchmarks by Language**

| Language | Oral Reading Fluency Benchmark |
|----------|-------------------------------|
| **Hindi** | 35 cwpm |
| **Marathi** | 40 cwpm |
| **Kannada** | 35 cwpm |
| **Oriya** | 30 cwpm |
| **English** | 30 cwpm |

## 7.6    Next Steps and Future Research

In the final session of the benchmarking workshop, participants summarized questions and comments they had about benchmarking procedures and the direction that future work could take in this area.

There was general agreement that much more work must be conducted to understand the development of reading in Indian languages. Indian languages have a transparent orthography, and questions were raised about the impact that has on the relationship between fluency and comprehension. The conversation also led to discussions about the need to better understand the role that visual complexity of languages plays in learning to read (particularly for the more visually complex South Indian languages). Some participants ultimately noted that ORF may be a poor predictor (and even poorer proxy) of reading comprehension in Indian languages and that alternative measures should be explored and examined. Related to the debate on using a single benchmark across languages, concerns were raised about the appropriateness of using a common benchmarking approach among the many dialects within a given language, as well as among students with different mother tongues who learn to read in the same language of instruction.

Several comments were also made about the use and interpretation of benchmarking data. For example, it was noted that standard setting is important for measuring outcomes but that it does not provide enough contextual information to be useful for teachers or school-level actors. Accordingly, it was recognized that the purpose of the activity was to set benchmarks for Standard 2. It was stressed by many that the use of benchmarks in the earliest grades (e.g., Standards 1 and 2) should be used to assess the system, as opposed to the child. Some participants raised concerns about assessing in the early grades at all, but it was generally agreed that some level of assessment is essential (in order to understand what is going well and where the challenges lie).

Lastly, there was significant discussion around the ability to reliably and validly measure comprehension for the purpose of benchmarking. The ASER tools do not directly assess comprehension; the EGRA comprehension assessment could be improved in many ways such as allowing children to read to the end of the passage and using separate, independent passages to assess fluency and comprehension. Given that research has not established whether fluency is a proxy for comprehension in Indian languages, it was suggested that these should all be seen as areas that are prime for future research. Participants noted that there is no shortage of data in India, but that it can sometimes be difficult to ensure that data are available and accessible.

Overall, the benchmarking activity was seen as a success by participants. As the first large-scale benchmarking activity across languages in India, it was noted that the exercise itself was incredibly valuable, but that there is still much to be learned. Therefore, beyond their use for reporting progress at endline, it is expected that these benchmarks will serve as the

starting point for a larger conversation regarding the appropriate methods for estimating and using EGR benchmarks in India.

# 8    Conclusion

The results of the initial 2017 data collection indicate a wide range (both within projects and across the projects) of ORF and reading comprehension scores, as well as a varied distribution of students reading across ASER's reading levels. The initial data collection has provided a standard common measure for all projects that can be used to report toward the Global Count and Indicator E.S.1-1. The initial assessment however is just the first step in being able to measure *progress* against these indicators. Progress will be calculated for each project following the final assessment. Project impact and gains in readings will be determined by comparing each project's end of year results to the initial assessment data (treatment and control). It is also important to remember that comparisons should not be drawn across projects. Each project is operating in different geographical regions, socio-economic contexts, and across languages. Furthermore, projects are in various stages of implementation, with some projects in their final year of implementation and some in their first year. These differences are evidenced in the data, as small to medium effect sizes are measured for some projects however it is not possible to attribute these differences to a treatment effect as balance between control and treatment groups cannot be assessed.

# 9    Next Steps

Data collection for the final assessment is planned for the end of February / beginning of March 2018. Data collection will be scheduled with each implementing partner to accommodate each state's holiday and end of year exam schedules. Implementing partners have been contacted and dates will be finalized for each project.

The purpose of the benchmarking workshop is to determine reading fluency benchmarks for each language. The process of setting benchmarks involved a wide group of regional and national stakeholders. Implementing partners, government stakeholders, and USAID/India were convened to ensure buy-in to the process and outcomes of the workshop. With early grade benchmarks set, existing data and endline data will be used to determine the percentage of students meeting each benchmark at baseline and endline. Furthermore, benchmarks can be used by implementing partners to inform future target setting.

# References

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

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.* [Online]. Retrieved from http://mhrd.gov.in/rte. [Accessed: December 25, 2017].

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. [Accessed: December 25, 2017].

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

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: ASER Centre. [Online]. Retrieved from http://www.asercentre.org/sampling/precision/reliability/validity/p/180.html. [Accessed: December 25, 2017].

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. [Online]. Retrieved from http://uis.unesco.org/sites/default/files/documents/understanding-what-works-in-oral-reading-assessments-2016-en.pdf. [Accessed: December 25, 2017].

# Annex A: Balance Testing of Initial Assessment EGRA Outcomes using Control Models

This appendix reports the results of an analysis of equivalence on Early Grade Reading Assessment (EGRA) key variables for the control and treatment groups by project for the initial assessment. Demonstrating equivalence or balance on key variables for the control and treatment groups at the initial assessment phase is important to ensure that factors other than the intervention do not account for differences in the outcomes.

**Why Baseline Equivalence Is Important.** In education, interventions are often evaluated by comparing baseline scores for treatment and control groups with their scores at the end of the intervention (gain scores). However, if the treatment and control groups differ in ways that are associated with the outcome at baseline (e.g., prior schooling or family wealth or education, characteristics of the teaching staff, etc.), then the treatment effect calculated at the endline might be partially due to these other factors. Lack of equivalence will result in mis-estimating the impact of the intervention.

As discussed in the executive summary, because this baseline was not conducted prior to the intervention, interpretation of a balance test is somewhat difficult. In addition, the students in the treatment group received program enhanced instructional support in Standard 1. Thus, we cannot disentangle the Standard 1 treatment effect from demographic differences between the control and treatment groups.

**How We Measure Baseline Equivalence.** The What Works Clearinghouse (WWC) specifies that a standardized mean difference between the control and treatment groups should be calculated to determine baseline equivalence (Institute of Educational Sciences, U.S. Department of Education, 2014). This standardized mean difference is calculated by dividing the difference in means between the baseline control and treatment groups by their pooled standard deviation, and is an effect size called Cohen's *d* (Cohen, 1988). The baseline equivalence is then determined by comparing the effect size to the values in *Figure A-1*.

**Figure A-1. WWC Standard for Baseline Equivalence**

| Absolute value of effect size ≤ 0.05 | 0.05 <Absolute value of effect size ≤ 0.25 | Absolute value of effect size >0.25 |
|---|---|---|
| Satisfies baseline equivalence | Statistical adjustment required to satisfy baseline equivalence | Does not satisfy baseline equivalence |

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

For effect sizes with absolute values less than or equal to .05, a difference in mean scores can be used to estimate the impact of the program. Values greater than .05 and less than or equal to .25 require "statistical adjustment," to adjust for observed differences on the particular characteristics. Difference models with added demographic covariates, based on the regression model above but with the addition of those variables requiring statistical adjustment, can be used to satisfy this requirement. This annex presents the findings of these regression models and their associated adjustment of the initial assessment balanced effect size. *Table A-1* shows the initial assessment mean differences and associated effect sizes from the control models. The two difference columns compare the control model and simple difference in means and effect size. The simple difference in means results are presented in *Table 4* in the executive summary. The control models show slight improvement in a lower difference in means and effect size. A successful model would reduce an effect size down by at least 0.1, or to a level less than 0.15. The only program for which the control model managed to achieve this aim is Start Early: Read in Time – Odisha, where the effect size went down from 0.34 to 0.18 for reading fluency and from 0.20 to 0.03 for reading comprehension.

**Table A-1.    Initial assessment equivalence testing with covariate models**

| Project-Location | Treatment | Oral Reading Fluency | | | Reading Comprehension % Correct | | |
|---|---|---|---|---|---|---|---|
| | | Control Model | | Difference in Means (Effect Size in Parentheses) | Control Model | | Difference in Means (Effect Size in Parentheses) |
| | | Estimates | Difference (Effect Size in Parentheses) | | Estimates | Difference (Effect Size in Parentheses) | |
| Nurturing Early Literacy – Maharashtra | Treatment | 30.8 | 4.2 (0.23) | 5.1 (0.28) | 41.7% | 1.9% (0.06) | 3.5% (0.12) |
| | Control | 35.0 | | | 43.6% | | |
| Start Early: Read in Time – Odisha | Treatment | 7.6 | 2.5 (0.18) | 5.0 (0.34) | 8.8% | 0.5% (0.03) | 3.8% (0.20) |
| | Control | 10.2 | | | 9.3% | | |
| Right to Read – Maharashtra | Treatment | 0.6 | 0.0 (0.01) | 0.1 (0.03) | 0.1% | 0.0% (0.02) | 0.1% (0.03) |
| | Control | 0.6 | | | 0.2% | | |
| Scaling Up Early Reading Intervention – Chhattisgarh | Treatment | 6.8 | 7.7 (0.56) | 8.0 (0.57) | 4.0% | 5.6% (0.37) | 5.9% (0.38) |
| | Control | 14.5 | | | 9.5% | | |
| Teacher Innovations in Practice – Uttar Pradesh | Treatment | 3.2 | 0.5 (0.05) | 0.6 (0.06) | 3.5% | -0.2% (-0.01) | -0.1% (0.01) |
| | Control | 3.7 | | | 3.3% | | |
| Scaling Up Early Reading Intervention – Uttarakhand | Treatment | 7.5 | 7.8 (0.49) | 8.4 (0.52) | 7.2% | 6.1% (-0.01) | 6.7% (0.34) |
| | Control | 15.3 | | | 13.3% | | |
| Nurturing Early Literacy – Rajasthan | Treatment | 1.4 | 0.3 (0.07) | 0.4 (0.07) | 0.8% | -0.1% (-0.03) | -0.1% (0.02) |
| | Control | 1.7 | | | 0.6% | | |
| Nurturing Early Literacy – Karnataka | Treatment | 2.2 | 0.4 (0.05) | 0.5 (0.07) | 1.1% | 0.7% (0.09) | 1.1% (0.13) |
| | Control | 2.5 | | | 1.8% | | |
| Start Early: Read in Time – Uttar Pradesh | Treatment | 2.6 | -0.3 (-0.03) | 0.0 (0.00) | 2.8% | -0.1% (-0.01) | 0.3% (0.02) |
| | Control | 2.3 | | | 2.7% | | |

The final two tables in this annex show the final covariate control models that were used to create the adjust effect sizes. ***Table A-2*** is for reading fluency by program, and ***Table A-3*** is for reading comprehension, by program.

**Table A-2.  Control models for oral reading fluency, by program**

| | Nurturing Early Literacy Maharashtra | Start Early: Read in Time Odisha | Right to Read Maharashtra | Scaling Up Early Reading Intervention Chhattisgarh | Teacher Innovations in Practice Uttar Pradesh | Scaling Up Early Reading Intervention Uttarakhand | Nurturing Early Literacy Rajasthan | Nurturing Early Literacy Karnataka | Start Early: Read in Time Uttar Pradesh |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | 4.188*** (3.84) | 2.535* (2.49) | 0.0348 (0.16) | 7.721*** (10.48) | 0.497 (0.78) | 7.846*** (8.84) | 0.317 (1.01) | 0.367 (1.03) | -0.259 (-0.38) |
| Wealth Index | 2.138*** (4.51) | 3.612*** (6.00) | 0.285** (2.78) | 0.589 (1.87) | | -0.920** (-3.01) | | 0.209 (1.03) | 0.729* (2.09) |
| Extra reading material at home | 3.231** (2.85) | 3.004* (2.51) | 0.312 (1.06) | 3.349*** (4.44) | | 3.541*** (4.00) | 0.505 (1.47) | 0.497 (1.18) | 2.056 (1.51) |
| Absent from school in last week | -4.577*** (-4.20) | | 0.466 (1.37) | -0.544 (-0.74) | -0.988 (-1.50) | -2.038* (-2.29) | -0.384 (-1.29) | -0.768* (-2.06) | 0.256 (0.23) |
| Someone read aloud at home | 1.331 (1.11) | 1.626 (1.38) | 0.306 (1.06) | | | | | | -0.443 (-0.48) |
| Child gender | 6.234*** (5.59) | 0.814 (0.73) | -0.148 (-0.52) | 1.214 (1.63) | | | | | 0.394 (0.37) |
| Help with homework at home | | 1.777 (1.29) | | | 0.529 (0.76) | | 0.629* (2.10) | -0.359 (-0.98) | 0.947 (1.12) |
| Child age | | | 0.678 (1.49) | 2.627*** (4.11) | 0.368 (0.94) | 0.768* (2.01) | 0.698* (2.00) | 0.953* (2.06) | -0.201 (-0.61) |
| Breakfast before school | | | -0.829 (-1.73) | -1.797 (-0.89) | 0.802 (0.84) | 2.505 (1.84) | | | 1.643* (2.54) |
| Electricity connection at home | | | | 1.593 (0.92) | 1.028 (1.59) | 4.872** (3.13) | 0.523 (1.87) | 0.607 (0.91) | 0.747 (0.79) |
| Constant | 16.16*** (7.44) | 7.061** (3.13) | -4.179 (-1.28) | -14.75** (-2.74) | -0.877 (-0.27) | -5.052 (-1.47) | -4.111 (-1.64) | -4.902 (-1.41) | 1.602 (0.38) |
| F p | 15.14 | 13.67 | 3.295 | 20.55 | 1.408 | 23.64 | 3.214 | 1.813 | 2.570 |
| N | 1359 | 840 | 1744 | 1707 | 1688 | 1629 | 1179 | 1664 | 1590 |

*t* statistics in parentheses
* *p* < 0.05, ** *p* < 0.01, *** *p* < 0.001

**Table A-3.  Control models for reading comprehension percent correct, by program**

| | Nurturing Early Literacy Maharashtra | Start Early: Read in Time Odisha | Right to Read Maharashtra | Scaling Up Early Reading Intervention Chhattisgarh | Teacher Innovations in Practice Uttar Pradesh | Scaling Up Early Reading Intervention Uttarakhand | Nurturing Early Literacy Rajasthan | Nurturing Early Literacy Karnataka | Start Early: Read in Time Uttar Pradesh |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | 1.854 (1.07) | 0.535 (0.39) | 0.0457 (0.32) | 5.561*** (6.90) | -0.179 (-0.23) | 6.072*** (5.55) | -0.128 (-0.45) | 0.687 (1.62) | -0.104 (-0.10) |
| Wealth Index | 3.231*** (4.14) | 3.894*** (5.32) | 0.0395 (1.05) | 0.392 (1.26) | | -0.562 (-1.48) | | -0.240 (-1.36) | 0.973* (1.98) |
| Extra reading material at home | 6.499*** (3.51) | 3.424* (2.39) | 0.169 (0.82) | 3.141*** (3.68) | | 3.530** (3.24) | 0.717* (2.11) | | 3.508 (1.82) |
| Absent from school in last week | -3.872* (-2.20) | | 0.397 (1.32) | -1.088 (-1.34) | -1.817* (-2.15) | -3.121** (-2.80) | 0.235 (0.85) | -0.609 (-1.38) | 0.947 (0.59) |
| Someone read aloud at home | 3.237 (1.70) | 1.479 (1.03) | 0.0764 (0.46) | | | | | | -0.205 (-0.16) |
| Child gender | 10.71*** (5.95) | 2.293 (1.63) | -0.256 (-1.26) | 0.559 (0.69) | | | | | 0.505 (0.36) |
| Help with | | 2.401 | | | 1.401 | | 0.296 | 0.267 | 1.837 |

| | Nurturing Early Literacy Maharashtra | Start Early: Read in Time Odisha | Right to Read Maharashtra | Scaling Up Early Reading Intervention Chhattisgarh | Teacher Innovations in Practice Uttar Pradesh | Scaling Up Early Reading Intervention Uttarakhand | Nurturing Early Literacy Rajasthan | Nurturing Early Literacy Karnataka | Start Early: Read in Time Uttar Pradesh |
|---|---|---|---|---|---|---|---|---|---|
| homework at home | | (1.63) | | | (1.76) | | (1.05) | (0.59) | (1.61) |
| Child age | | | 0.704 (1.02) | 2.616*** (3.75) | | 0.656 (1.39) | 0.542 (1.50) | 0.767* (2.03) | |
| Breakfast before school | | | 0.164 (1.33) | -1.338 (-0.78) | 1.960* (2.34) | 1.401 (0.81) | | | 1.401 (1.83) |
| Electricity connection at home | | | | | 1.479 (1.82) | 5.800** (2.90) | 0.634** (2.65) | 1.209 (1.94) | 0.977 (0.69) |
| Constant | 14.67*** (4.15) | 5.673* (2.01) | -4.942 (-1.00) | -14.96** (-2.87) | 0.736 (0.62) | -4.006 (-0.93) | -3.989 (-1.51) | -5.719* (-2.01) | -0.884 (-0.24) |
| F p | 12.82 | 9.381 | 1.195 | 12.55 | 2.746 | 12.50 | 2.196 | 1.937 | 3.378 |
| N | 1441 | 851 | 1759 | 1724 | 1723 | 1635 | 1180 | 1718 | 1618 |

*t* statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Annex B: Overview of USAID/India Early Grade Reading 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 learners, 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 by developing and disseminating teaching/learning materials that are contextualized to meet the diverse needs of the students. The key strategies are to build on the child's 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 early grade reading outcomes in the states of Delhi and Uttar Pradesh by positively impacting the teaching practices of 14,657 teachers and the early grade reading 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 | Right to Read | English Helper Education Technologies Private Limited | September 2015 | September 2019 | Maharashtra, West Bengal, Gujarat, Delhi, Tamil Nadu, Telangana, Karnataka, Punjab |
|---|---|---|---|---|---|

**Description:** This project addresses the shortage of English language teachers and poor English language skills among students in government-run primary schools. English Helper 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 about 5,000 primary schools.

**Scalability:** Since the launch of the project and subsequent visibility as a result of the USAID partnership, Right to Read 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 ICT 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 4 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 learner 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: District Sampling Notes

During project planning decisions were made about which languages and districts to include in the evaluation sample plan. *Table C-1* documents which languages and districts were, or were not, selected for each project.

**Table C-1.    District sampling decisions**

| Project | Location | District Sampling Decisions |
|---|---|---|
| School Excellence Program | Gujarat | The assessment proposes to cover only Gujarati medium schools. Inclusion of additional languages would have required the sample size to increase by 60 treatment schools per language. Due to challenges in obtaining permission to test in schools, data collection was not conducted for this project. |
| Start Early: Read in Time | Uttar Pradesh and Odisha | For the CARE program in Uttar Pradesh, 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, with separate treatment and control samples for each state. |
| Teacher Innovations in Practice | Uttar Pradesh | All 10 districts in Uttar Pradesh will be included. Two of these districts, Varanasi and Rae Bareli, are also covered in the Annual Status of Education Report's (ASER's) independent evaluation of the program. |
| Right to Read | Maharashtra and West Bengal | Schools being assessed for Right to Read are extension schools only; meaning 2017–2018 will be the first year of implementation. For logistical considerations, schools in 6 districts will be included in the Maharashtra assessment. Although training of assessors was completed for West Bengal in January 2018, permissions were not granted to collect data from the local government. |
| Nurturing Early Literacy | Rajasthan, Karnataka, and Maharashtra | Nurturing Early Literacy is being implemented by 2 separate organizations (essentially 2 different interventions; one operated by Bodh Shiksha Samiti and the other by Room to Read). The schools under Bodh's intervention will be included in this assessment as Room to Read is already included in this study under the Scaling Up Early Reading Intervention project. |
| Scaling Up Early Reading Intervention | Uttarakhand and Chhattisgarh | Separate treatment and control samples for each state. |

Each sampled treatment school was matched with a control school from the same block (within the same district) based on following criteria: school management type, Standard present in school (primary only – Standard 1 to 5, primary with upper primary – Standard 6 to 8, and so on), enrollment in Standard 2, number of teachers appointed, and availability of computers and libraries for children's use.

*Table C-2* shows the number of control schools, by project location, where exceptions were made to the above matching process during sampling for baseline data collection.

**Table C-2.  Number of control schools by project locations that are not matched to treatment schools**

| S. No. | Name of the program | Name of District | Number of schools where: | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | | District of control does not match treatment | Block of control does not match treatment | School management of control does not match treatment | School category of control does not match treatment |
| 1 | Start Early: Read in Time - Uttar Pradesh | Balrampur | 0 | 0 | 0 | 0 |
| | | Bahraich | 0 | 0 | 0 | 0 |
| | | Shrawasti | 0 | 0 | 0 | 0 |
| | | Hardoi | 0 | 0 | 0 | 0 |
| | | Gonda | 0 | 0 | 0 | 0 |
| 2 | Start Early: Read in Time - Odisha | Mayurbhanj | 0 | 0 | 0 | 7 |
| | | Keonjhar | 0 | 0 | 0 | 0 |
| | | Dhenkanal | 0 | 0 | 0 | 0 |
| 3 | Teacher Innovations in Practice - Uttar Pradesh | Faizabad | 0 | 0 | 0 | 0 |
| | | Varanasi | 0 | 0 | 0 | 0 |
| | | Unnao | 0 | 0 | 0 | 0 |
| | | Barabanki | 0 | 0 | 0 | 0 |
| | | Kanpur Nagar | 0 | 0 | 0 | 0 |
| | | Chandauli | 0 | 0 | 0 | 0 |
| | | Mirzapur | 0 | 0 | 0 | 0 |
| | | Lucknow | 0 | 0 | 0 | 0 |
| | | Rae Bareli | 0 | 0 | 0 | 0 |
| | | Jaunpur | 0 | 0 | 0 | 0 |
| 4 | Right to Read - Maharashtra | Latur | 0 | 0 | 5 | 13 |
| | | Pune | 0 | 0 | 7 | 9 |
| | | Oshmanabad | 1 | 1 | 4 | 8 |
| | | Jalgaon | 0 | 0 | 5 | 5 |
| | | Nagpur | 0 | 16 | 18 | 20 |
| | | Solapur | 0 | 0 | 7 | 7 |
| 5 | Scaling up Early Reading Intervention - Chhattisgarh | Raipur | 0 | 0 | 0 | 0 |
| | | Baloda Bazar | 0 | 0 | 0 | 0 |
| 6 | Scaling up Early Reading Intervention - Uttarakhand | Almora | 53 | 53 | 0 | 0 |
| | | Dehradun | 0 | 0 | 0 | 0 |
| | | Udham Singh Nagar | 0 | 0 | 0 | 0 |

| S. No. | Name of the program | Name of District | Number of schools where: | | | |
|---|---|---|---|---|---|---|
| | | | District of control does not match treatment | Block of control does not match treatment | School management of control does not match treatment | School category of control does not match treatment |
| 7 | Nurturing Early Literacy - Rajasthan | Sirohi | 0 | 0 | 0 | 0 |
| 8 | Nurturing Early Literacy - Karnataka | Yadagiri | 0 | 0 | 0 | 0 |
| 9 | Nurturing Early Literacy - Maharashtra | Satara | 0 | 0 | 0 | 0 |

# Annex D: Enrollment and Attendance in Surveyed Schools

Within each sampled school, 20 students (10 boys and 10 girls) were to be sampled. Hence, in each project location, 2,400 students (120 schools x 20 students) were expected to be sampled. Care was taken to schedule the actual data collection at a time when schools had finalized their enrollment registers for the year and classes were settled. However, due to low enrollment and attendance on the day of visit, the target sample of students was not met for each project location (see *Table D-1*).

**Table D-1.  Number of schools with enrollment and attendance of fewer than 20 students, by project location**

| Sno. | Program | Number of Schools Surveyed | | | Number of Schools with Enrollment Fewer Than 20 Students | Number of Schools with Attendance Fewer Than 20 Students |
|---|---|---|---|---|---|---|
| | | Treatment | Control | Total | | |
| 1 | Scaling Up Early Learning Intervention – Uttarakhand | 90 | 90 | 180 | 130 | 147 |
| 2 | Scaling Up Early Learning Intervention – Chhattisgarh | 60 | 60 | 120 | 44 | 69 |
| 3 | Nurturing Early Literacy – Rajasthan | 60 | 60 | 120 | 59 | 110 |
| 4 | Teacher Innovations in Practice – Uttar Pradesh | 70 | 70 | 140 | 46 | 113 |
| 5 | Start Early: Read in Time – Uttar Pradesh | 70 | 70 | 140 | 32 | 112 |
| 6 | Start Early: Read in Time – Odisha* | 60 | 60 | 120 | 105 | 118 |
| 7 | Nurturing Early Literacy – Karnataka | 60 | 60 | 120 | 42 | 68 |
| 8 | Nurturing Early Literacy – Maharashtra | 70 | 70 | 140 | 113 | 122 |
| 9 | Right to Read – Maharashtra** | 67 | 44 | 111 | 24 | 50 |
| 10 | Right to Read – West Bengal | Data collection could not begin due to lack of permission from the local government. | | | | |
| 11 | School Excellence Program – Gujarat | Data collection could not begin due to lack of permission from the local government. | | | | |
| | Total | 607 | 584 | 1191 | 595 | 909 |

Additionally, data collection could not be completed in two project locations:
*Start Early Read in Time – Odisha: field teams were not allowed to carry out data collection in 2 (out of 3 districts) – Dhenkanal and Keonjhar.
**Right to Read – Maharashtra: permission was being arranged school by school and not a blanket permission. Hence, many schools denied permission to collect data. As a result, many control schools in Pune and Latur districts were not surveyed.

# Annex E: 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 have not been included here, with consideration of document size, but are available upon request.

**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
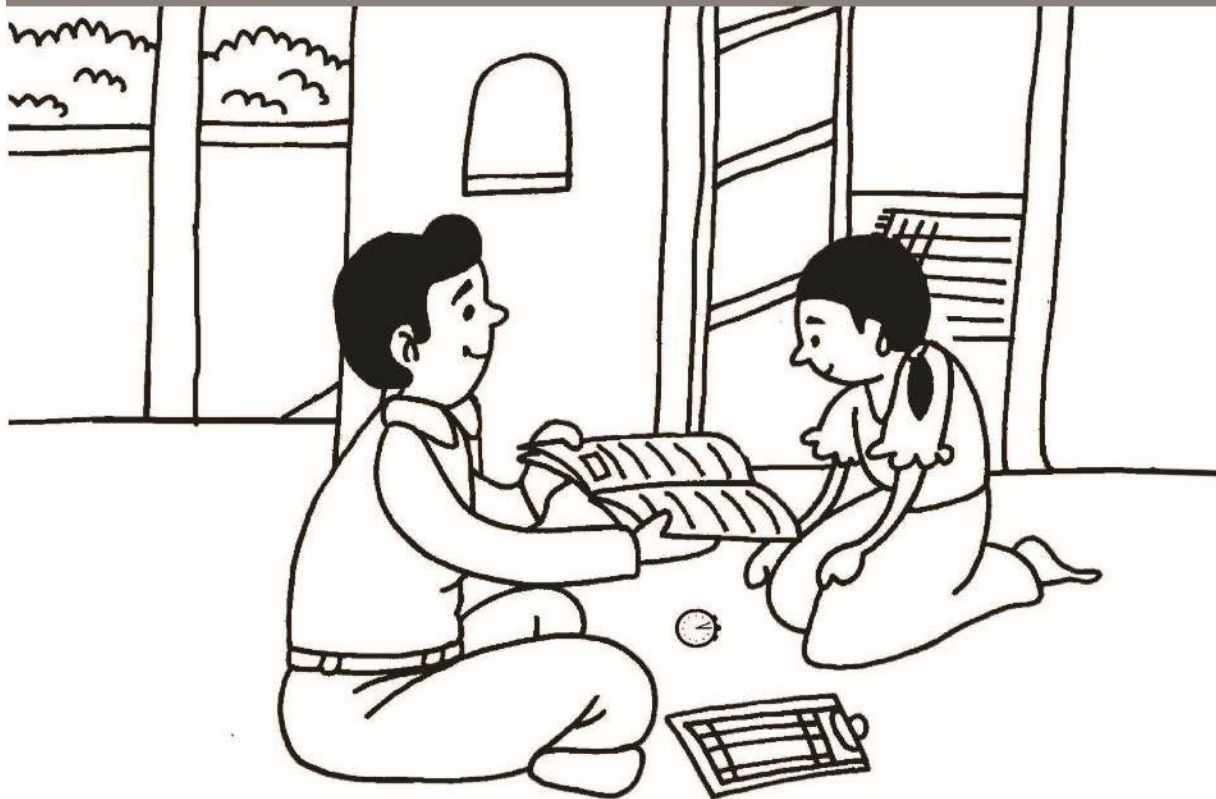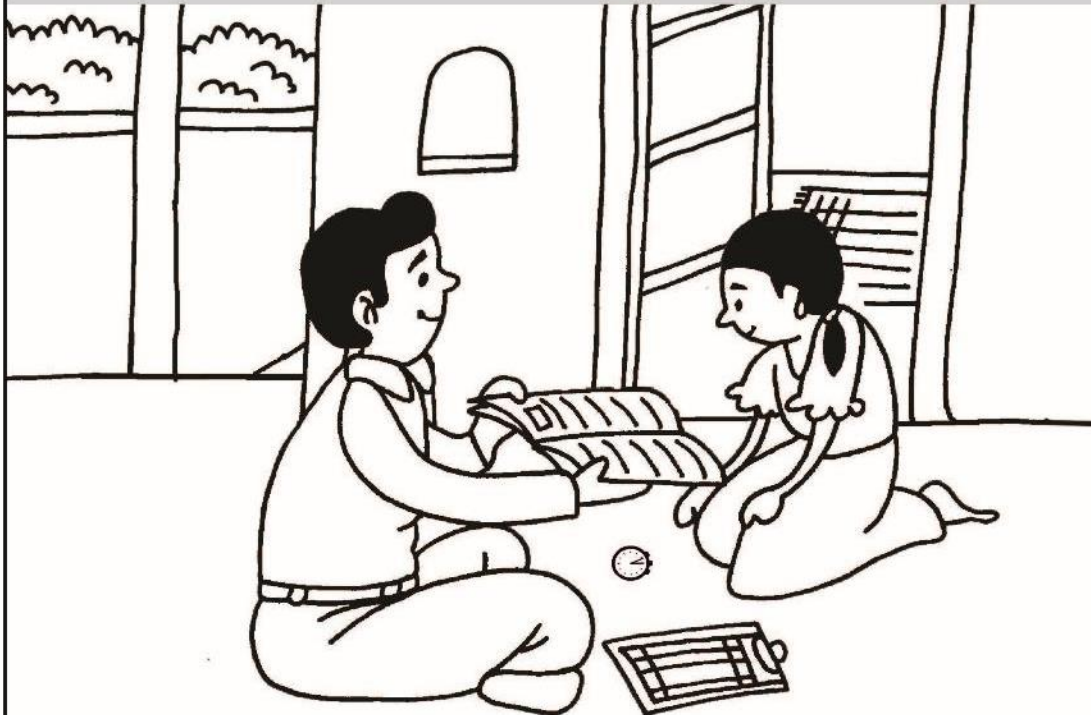


For surveyor – don't give this booklet to the child.

## Survey Information

| Type (Pick One) | Surveyor | Monitor | |
|---|---|---|---|
| State name | | SCHOOL CODE | |
| School name | | | |
| Surveyor ID | | Date of survey | |
| Surveyor name | | Mobile number | |

**ASER CENTRE, ASSESSMENT UNIT**　　　　　　　　**USAID - BASELINE VISIT SEPTEMBER 2017**

## 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 <u>ONLY</u> 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)

- <u>**This is NOT a test and it will not affect your grade at school.**</u>

- 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?

*Analysis of Early Grade Reading Assessment in India*

## 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 | Comprehension |
|---|---|
| *Surveyor tells the child:* "Read this story aloud but carefully. When you finish, I will ask you some questions about what you have read. <br> 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. <br> Put your finger on the first word of the story. Ready? Begin." | *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." |

| Reading | Comprehension |
|---|---|
| • Start the timer as soon as you say "Begin". <br> • When the timer reaches 0, tell the child to "Stop". <br> • If a child hesitates while reading or stops on a word for 3 SECONDS, say "Go on." <br> • 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. <br> (/) Mark any incorrect or unread words with a slash (/). <br> (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. <br> (]) Mark the final word read with a bracket (]) at the end of 60 seconds. | After the child is finished reading, REMOVE the story from in front of the child. <br> • 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. <br> • 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 (*). <br> • 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. <br> • Do not ask the question more than twice. <br> • If a child says "I don't know," mark the child at 'incorrect'. |

| Story for reading <br> *(Child reads out of his/her booklet)* | Number of correct words | Question <br> [Answer] | N.A. | Correct | Incorrect | No response |
|---|---|---|---|---|---|---|
| Gita goes to school every day*. Her mother gave her a book. <br> 1  2  3    4    5   6    7      8      9    10 11  12 | ☐/12 | Where does Gita go every day? <br> [School] | | | | |
| The book had a red cover. Gita read the book every morning on her <br> 13  14    15 16 17    18    19   20   21 22    23    24    25 26 <br> way to school*. <br> 27  28  29 | ☐/17 | What did Gita do on her way to school? <br> [Read] or [ Read the book] | | | | |
| Gita learned many new words from the book*. <br> 30    31    32    33  34    35   36  37 | ☐/8 | What did Gita learn from the book? <br> [Words] or [ New words] | | | | |
| That made her teacher happy. The teacher gave Gita another book*. <br> 38   39   40    41    42   43   44    45   46   47    48 | ☐/11 | Who gave another book to Gita? <br> [Teacher] | | | | |
| It had nice stories. She showed it to all her friends*. <br> 49 50  51    52    53    54    55 56 57 58    59 | ☐/11 | Why did Gita show the book to all her friends? <br> [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: ☐

## Part 2 - Child Questionnaire

*(Show pictures to the child for Q8, Q9 and Q10.)*

**Q8.** Generally, where does drinking water come from at your home?
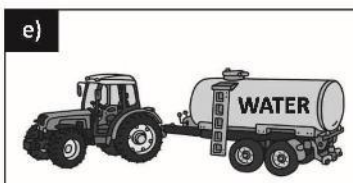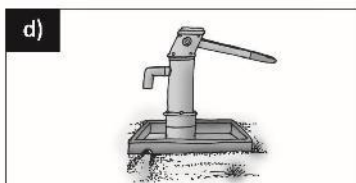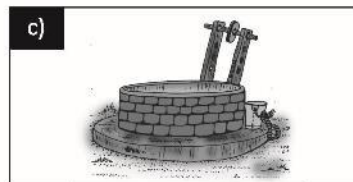


**Q9.** Generally, how is food cooked at your home?
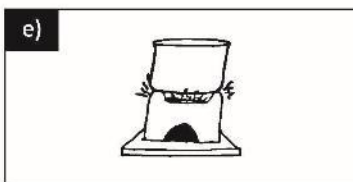


**Q10.** Do you have the following items in your home?
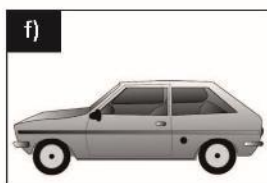
# Annex F: Pilot Details and Full Assessment Psychometrics by Language

A pilot was conducted between August 29 and September 2, 2017. The pilot was completed by experienced Annual Status of Education Report (ASER) data collectors in 8 states and 8 districts. Approximately 2,400 students in Standard 2 were assessed. *Table F-1* gives details of the pilot state by state, indicating the districts, languages, and number of students assessed in each location.

**Table F-1.    Sample of students for pilot data collection by state and district**

| Pilot States | Pilot Districts | Language | Number of Students |
|---|---|---|---|
| Chhattisgarh | Mahasamund | Hindi | 78 |
| Gujarat | Gandhinagar, Mehsana | Gujarati | 212 |
| Karnataka | Mysore | Kannada | 372 |
| Odisha | Puri | Oriya | 410 |
| Maharashtra | Satara | Marathi | 366 |
| Maharashtra | Satara | English | 303 |
| Uttar Pradesh | Varanasi | Hindi | 628 |
| Uttarakhand | Dehradun | Hindi | 41 |
| Total | | | 2410 |

The purpose of the pilot was to collect qualitative information on the following administration questions:

- Does the order of testing, i.e., administering ASER followed by oral reading fluency (ORF) or ORF followed by ASER, have any significant impact on children's test performance?

- Should 1 or 2 data collectors be deployed to schools for data collection?

Data were also collected to assess the reliability and validity of the assessments and address the following questions:

- Can raters reliably note down reading errors on ORF and accurately rate responses to the reading comprehension questions?

- Does the reading comprehension component demonstrate good reliability as indexed by internal consistency?

- What is the evidence for concurrent validity of the ASER, ORF, and reading comprehension assessments?

**Pilot Results**

This annex provides a summary of the pilot study results.

*Does the order of testing, i.e., administering ASER followed by ORF or ORF followed by ASER, have any significant impact on children's test performance?*

About half the pilot study sample for each language administered the ASER assessment followed by the ORF and reading comprehension assessments. The order of test administration was reversed for the other half of the sample, i.e., ORF and reading comprehension assessments were administered first, followed by the ASER assessment. Given that the two halves of the sample are independent groups, we cannot assign differences between the two groups' performance on either assessment to testing order effects. It is possible that the ability composition of the two groups was different. One limited way to evaluate test performance for the two groups is to estimate the difference in the magnitude of association between ASER and ORF for the two groups. The assumption here is that the association between ASER and ORF should be similar (not significantly different) for the two sub-samples. In addition, we also conducted a difference in mean analysis to understand whether the reading ability levels and fluency rates were similar or not across the two sub-samples.

For all languages, the difference in mean estimates were not significant, thus indicating that the reading ability levels on ASER were similar across the two sub-samples and that fluency rates as well were similar across the two sub-samples. The magnitude of the correlation coefficients between ASER and ORF were not significantly different across the sub-samples for each language. These results suggest no test order effects. Individual pilot reports for each language can be provided upon request.

Given the absence of test order effects and based on the feedback from the examiners, it was determined that the ASER assessment would be administered first, followed by the ORF and reading comprehension questions. Examiners and monitors noted that students seemed intimidated by having to begin with reading the ORF passage. Instead, the adaptive nature of the ASER assessment that included letter and word reading levels as well as a Standard 1-level passage allowed students to warm up to the more challenging task of reading passages and responding to the corresponding comprehension questions.

*Should 1 or 2 data collectors be deployed to schools for data collection?*

Use of two data collectors is the typical administration modality for the ASER assessment when in households. During the pilot, in some schools two data collectors were deployed, and in other schools one data collector was deployed. However, with the addition of the ORF component, data collectors reported back from the field that it was difficult to administer the assessment with two data collectors and two stopwatches. Hence, based on data collector feedback and monitor observations, one data collector was found to be sufficient to administer the ASER and the modified EGRA assessments.

*Can raters reliably score student's errors on the ORF reading passage and reliably mark responses to the reading comprehension questions?*

Reading comprehension: This assessment was comprised of 5 questions, however the number of questions administered to each child differed, as it was contingent on the amount of text read in the span of one minute. Hence, Cohen's kappa as an index of inter-rater reliability was estimated for each individual question as well as for all questions together. Cohen's kappa estimates ranged from 0.88 to 1 for all languages (except English), suggesting high inter-rater reliability for the reading comprehension assessment. The data for English were insufficient, with a maximum of 9 respondents.

ORF: Inter-rater reliability for reading errors (or number of words correctly read) was evaluated on the basis of a correlation coefficient and the mean absolute error. The correlation coefficients indicate whether the relative rankings of children's errors were similar for the two raters and the mean absolute error indicates whether the difference in absolute values was similar (see pilot reports for details). The correlation coefficients for all languages ranged from .96 to .99, and the

difference between raters' and monitors' mean absolute error is zero for all languages. Hence, the evidence suggests high inter-rater reliability for scoring errors on the ORF reading passages.

Inter-rater reliability was high across all languages, although a few language groups had small samples.

*Does the reading comprehension component demonstrate good reliability as indexed by internal consistency?*

Reliability based on internal consistency for the reading comprehension task ranged from "adequate" for Kannada (Cronbach's alpha = 0.77) and Marathi (Cronbach's alpha = 0.71) to "good" for Gujarati (Cronbach's alpha = 0.8). Reliability based on internal consistency for the reading comprehension tasks in English could not be estimated as substantially high numbers of non-readers in the pilot sample resulted in insufficient data. Reliability for Oriya (Cronbach's alpha = 0.6) was found to be less than adequate.

*How does reading performance on ASER relate to reading performance on ORF, and how does reading performance on both assessments relate to children's reading comprehension?*

The magnitude of the associations between the reading assessments and each reading assessment with the comprehension component ranged from moderate to strong ($r = 0.5$ to 0.9). The associations for English are mitigated ($r = 0.4$ and 0.6) due to high floor effects given the substantial number of non-readers. In addition, the mapping of performance on ORF with the different ASER reading levels suggests a high level of consistency. These results corroborate the research findings from an earlier evaluation of children's Hindi reading performance on ASER and EGRA (Vagh, 2016). In sum, these results provide favorable evidence for the validity of the ASER assessment, the ORF assessment, and reading comprehension assessments.

Studies indicate that the ASER assessments provide reliable and valid information about children's early reading abilities (Vagh, 2009; 2016).

**Initial Assessment Results**

Psychometric analyses of the assessments, notably reliability based on internal consistency for the reading comprehension assessment, and concurrent validity based on performance on the ASER, ORF, and reading comprehension assessments, were replicated for the baseline samples for each of the five languages. In addition, inter-rater reliability for the ASER assessment and the ORF assessment (the recording of reading errors) were evaluated, as monitors independently scored children's reading performance for a sub-sample while observing the data collectors. These results by language are presented in *Table F-2*.

We address the following questions based on the data collected for the baseline evaluation:

- Can raters reliably score a student's errors on the ORF reading passage, and can raters reliably rank children at the different reading levels on the ASER assessment?

- Does the reading comprehension component demonstrate good reliability as indexed by internal consistency? And how do these findings compare to the pilot study findings?

- What is the evidence for concurrent validity of the ASER, ORF, and reading comprehension assessments? And how do these findings compare to the pilot study findings?

*Can raters reliably score a student's errors on the ORF reading passage, and can raters reliably rank children at the different reading levels on the ASER assessment?*

ORF: Inter-rater reliability for reading errors (or number of words correctly read) was evaluated on the basis of a correlation coefficient and the mean absolute error. The correlation coefficients indicate whether the relative rankings of children's errors were similar for the two raters, and the mean absolute error indicates whether the difference in absolute values was similar. (Individual language reports can be provided upon request.) The correlation coefficients for all languages ranged from 0.86 to 0.97, and the difference in mean between data collectors' and monitors' mean absolute error was very low, ranging from 0 to 2. Hence, the evidence suggests high inter-rater reliability for scoring errors on the ORF reading passages.

ASER: Cohen's kappa estimates for the five languages ranged from 0.86 to 0.97, indicative of high inter-rater reliability. However, the data for some of the reading level categories was sparse on the ASER assessment.

*Does the reading comprehension assessment demonstrate good reliability as indexed by internal consistency? And how do these findings compare to the pilot study findings?*

Reliability based on internal consistency was adequate for Hindi (Cronbach's alpha=0.71) and Marathi (Cronbach's alpha=0.68), and these results were similar to those reported for the pilot.

Reliability for English was also adequate (Cronbach's alpha = 0.76); however, the sample size was small due to low reading attainments. About 93% of the children who were administered the English assessments were non-text readers. We were unable to evaluate reliability of the reading comprehension assessment in the pilot study due to low reading attainment.

For Kannada, Question 1 seems to have performed poorly, and the reliability excluding the first question improved from a Cronbach's alpha estimate of 0.59 to 0.73. Note that Question 1 was not flagged as a poorly performing item in the pilot study. However, the reading attainment levels for the baseline sample were lower compared to the pilot sample. It is plausible that the item was performing poorly for this specific population given their lower reading attainments. Excluding the first question from all further analyses with the Kannada reading comprehension assessment seems appropriate given the substantially higher reliability without this question. Alternatively, all further analyses can be conducted with and without the first question to ensure the robustness of the findings.

The reliability estimates for Oriya were less than adequate, and these findings were similar to the pilot study findings. Here too, we note extremely low reading attainment. The group of children who were administered the comprehension questions reported reading just 19 words correct per minute on average. For Oriya as well, it is entirely plausible that a constricted range of performance given the low reading attainments hampers the robustness of the assessment tool. Given the lack of substitute items, all findings based on the Oriya reading comprehension assessment should be interpreted with some caution.

*What is the evidence for concurrent validity of the ASER, ORF, and reading comprehension assessments? And how do these findings compare to the pilot study findings?*

The magnitude of the associations between ASER and ORF was strong ($r$s ranged from 0.7 to 0.85) with the exception of English ($r = 0.41$), and the magnitude of the associations between the reading comprehension assessments and the reading assessments ranged from low to moderate to high ($r$s range from 0.35 to 0.86). The correlation coefficients were most likely attenuated by the restricted range in performance due to low reading attainments. The mapping of performance on ORF with the different ASER reading levels provides additional and favorable evidence for the validity of the reading assessments. These findings are similar to the pilot study samples.

**Table F-2.    Summary of psychometric analyses: Pilot study and baseline assessment**

| Language | n | Order of Test Administration[11] | Inter-rater Reliability | Reliability of the Reading Comprehension Component | Validity |
|---|---|---|---|---|---|
| **Hindi (pilot sample)** | 744 | Results indicate that Annual Status of Education Report (ASER) and oral reading fluency (ORF) are strongly correlated and the magnitudes of the correlation coefficients for Group A (n=382) are not significantly different from the coefficients of Group B (n=362).<br><br>We conclude there are no test order effects on student performance. | Reading Comprehension (Q1 n=55, Q2 n=35, Q3 n=25, Q4 n=21, Q5 n=9) and ORF (n=105):<br><br>The sample size for evaluating inter-rater reliability for the comprehension questions is small. Some of the cell values in the cross tabular distributions are less than 5. Therefore, we cautiously conclude that a single examiner is able to reliably record errors ($r = 0.99$) and accurately rate student responses to the comprehension questions (Cohen's kappa estimates range from 0.94 to 1 for Q1 to Q4). The data are insufficient to estimate Cohen's kappa for Q5. | The inter-item correlations between all the comprehension questions range from moderate (0.6) to strong (0.8). The overall Cronbach's alpha coefficient of 0.71 (n=244) is acceptable. Typically, reliability coefficients of 0.70 and higher are considered adequate.<br><br>Reliability based on internal consistency for the reading comprehension task is fairly good. | The concurrent validity coefficients between the ASER and the ORF assessment is high ($r = 0.80$, p<.05, n=744). The associations of the reading comprehension assessments with the reading assessments are also high ($r = 0.80$ and 0.90, p<.05, n=2717). In addition, the average fluency rates progressively increase (from 0 to 47 correct words per minute [cwpm]) with higher reading levels on the ASER assessment. These results provide favorable evidence for the validity of the ASER assessment, the ORF assessment, and reading comprehension assessments. |
| **Hindi (baseline sample)** | 8,347 | Not applicable | ASER (n=2584) and ORF (n=2588):<br><br>Inter-rater reliability for the ASER reading assessment is high (Cohen's kappa coefficient = 0.93). However, given the limited variation in reading levels and the sparse data in some categories, we cautiously conclude that the inter-rater reliability for the ASER reading assessment is good.<br><br>The inter-rater reliability between monitor and examiner | The Cronbach's alpha coefficient for the reading comprehension component is 0.71 (n=1337), which is adequate. This is identical to the reliability estimate noted for the pilot sample. | The concurrent validity coefficients between the ASER and the ORF assessment are fairly high ($r = 0.78$, p<.05, n=8347), while the associations of the reading assessments with the reading comprehension assessment are in the moderately strong range ($r = 0.63$ and 0.75, p<.05, n=2717). In addition, the average fluency rates progressively increase (from 0 to 39 cwpm) with higher reading levels on the ASER |

---

[11] From hereon we combined Groups A and B. Where appropriate, we did analyze for Group A and Group B independently. The statistics for Groups A and B were found to be similar to the full sample. Hence, these additional analyses are not reported in the current report but are available on request.

*Analysis of Early Grade Reading Assessment in India*

| Language | n | Order of Test Administration[11] | Inter-rater Reliability | Reliability of the Reading Comprehension Component | Validity |
|---|---|---|---|---|---|
| | | | for noting down words read correctly on the ORF assessment is high (Pearson's correlation coefficient = 0.98). The mean absolute error between the examiners' and monitors' reporting of words read correctly is zero words. This suggests a high degree of inter-rater reliability for the reporting of words read correctly. | | assessment. These high levels of mapping across the two reading assessments provide additional and favorable evidence for concurrent validity. These findings are similar to the trends noted for the pilot study sample. |
| **Gujarati** | 212 | Results indicate that ASER and ORF are strongly correlated and the magnitudes of the correlation coefficients for Group A (n= 108) are not significantly different from the coefficients for Group B (n= 104). We conclude there are no test order effects on student performance. | Reading Comprehension (Q1 n=135, Q2 n=97, Q3 n=23, Q4 n=4, Q5 n=0) and ORF (n=131): The results indicate that a single examiner is able to reliably record errors and accurately rate student responses to the comprehension questions. There are insufficient data for the fourth fact retrieval question and the fifth inferential question. Despite insufficient data, we recommend retaining the fourth and fifth comprehension questions and if possible using data from the main project to evaluate these items. | Estimates indicate positive and strong item-test correlation coefficients and an overall Cronbach's alpha value of 0.8 (n=135). Even though Q5 was not administered to any students and cannot be adequately evaluated, we suggest retaining the inferential question. Data from the main study can help evaluate this item and determine whether it should be retained in follow-up assessments. | The concurrent validity coefficients between the ASER and the ORF assessment are strong ($r$ = 0.90). The associations of the reading comprehension assessments with the reading assessments are moderately low (ASER: $r$ = 0.50 and ORF: $r$ = 0.50). These results provide some favorable evidence for the validity of the ASER assessment, the ORF assessment, and reading comprehension assessments. |
| **English (pilot sample)** | 303 | Results indicate that ASER and ORF are strongly correlated and the magnitudes of the correlation coefficients for Group A (n= 167) are not significantly different from the coefficients for Group B (n= 136). | Reading Comprehension (Q1 n=17, Q2 n=4, Q3 n=3, Q4 n=0, Q5 n=0) and ORF (n=181): The sample size for evaluating inter-rater reliability for the comprehension questions is small, as only 12% of this sample are readers. Cohen's | Reliability based on internal consistency for the reading comprehension task cannot be estimated due to substantially high numbers of non-readers in the pilot sample resulting in insufficient data. | The concurrent validity coefficients between the ASER and the ORF assessment are low ($r$ = 0.40 n=303). The associations of the reading comprehension assessments with the reading assessments are moderately low ($r$ = 0.60 |

| Language | n | Order of Test Administration[11] | Inter-rater Reliability | Reliability of the Reading Comprehension Component | Validity |
|---|---|---|---|---|---|
| | | We conclude there are no test order effects on student performance. | kappa for the first comprehension question is a "perfect" 1, however this is based on just 17 student responses. There are insufficient or no data to evaluate the remaining comprehension questions.<br><br>Inter-rater reliability for the recording reading errors is high ($r = 0.99$). | Analysis of the data from the main study can help evaluate reliability of the reading comprehension component. | and 0.60, p<.05, n=29). High floor effects due to low reading attainments potentially attenuate the magnitude of the correlation coefficients.<br><br>The average fluency rates progressively increase (from 0 to 33 cwpm) with higher reading levels on the ASER assessment. These results provide some favorable evidence for the validity of the ASER assessment, the ORF assessment, and reading comprehension assessments. |
| **English (baseline sample)** | 1,833 | Not applicable | ASER (n=712) and ORF (n=712):<br><br>Inter-rater reliability for the ASER reading assessment is high (Cohen's kappa coefficient = 0.96). However, given the limited variation in reading levels and the sparse data in some categories, we cautiously conclude that the inter-rater reliability for the ASER reading assessment is high.<br><br>The inter-rater reliability between the examiner and monitor for noting down words read correctly on the ORF assessment is high (Pearson's correlation coefficient = 0.98). The mean absolute error between the examiners' and monitors' reporting of words read correctly is less than a word (MAE = 0.05). This suggests a high degree of inter- | The Cronbach's alpha coefficient for the reading comprehension component is 0.76 (n=133), which is adequate. However, these results need to be interpreted with some caution since the number of students who were administered Q3 (n=18), Q4 (n=11), and Q5 is small.<br><br>Reliability of the reading comprehension assessment could not be assessed in the pilot study due to insufficient data (high percentage of nonreaders). However, in the baseline sample as well, despite a total sample of 1,833 students, low reading attainments impact the amount of data available to evaluate the reading comprehension assessment. | The concurrent validity coefficients between the ORF and ASER are low-moderate ($r$ = .41, n=1833) and between the reading comprehension and reading assessments range from moderate to low-moderate ($r = 0.54$ and $r = 0.48$, n=133). These coefficients are most likely attenuated given the presence of high floor effects. However, a high level of mapping is noted across the two reading assessments with the average fluency rate increasing with each higher reading level on ASER (M=0 for beginners and students at the letter level; M=6 for students at the word level; M=25 for students at Standard 1 text level; and M=54 for students at the Standard 2 text level). In addition, 92% and 93% of the sample are classified as non- |

| Language | n | Order of Test Administration[11] | Inter-rater Reliability | Reliability of the Reading Comprehension Component | Validity |
|---|---|---|---|---|---|
| | | | rater reliability for the reporting of words read correctly. | | text readers on the ORF and the ASER assessments. These findings provide favorable evidence for concurrent validity. |
| **Kannada (pilot sample)** | 372 | Results indicate that ASER and ORF are strongly correlated and the magnitudes of the correlation coefficients for Group A (n=202) are not significantly different from the coefficients of Group B (n=170).<br><br>We conclude there are no test order effects on student performance. | Reading Comprehension (Q1 n=109, Q2 n=70, Q3 n=44, Q4 n=31, Q5 n=26) and ORF (n=208):<br><br>The inter-rater reliability estimates for reading errors ($r$ = 0.99, n=208) and the comprehension questions (Cohen's kappa estimate ranges from 0.93 to 1); The number of students administered the successive comprehension questions decreases, and the sample sizes for comprehension questions 4 and 5 are relatively smaller. However, these estimates suggest that there is high agreement (and low ambiguity) when recording errors or evaluating student responses to the comprehension questions. | The inter-item correlations between all the comprehension questions range from moderate (0.6) to strong (0.8). The overall Cronbach's alpha coefficient for the reading comprehension component is 0.77 (n=182). Typically, reliability coefficients of 0.70 and higher are considered adequate.<br><br>The reading comprehension questions appear to be performing well, and reliability based on internal-consistency is adequate. | The concurrent validity coefficient of ASER and ORF is high ($r$ =0.85) and of the reading comprehension assessment with ASER and ORF is high as well ($r$ =0.76 and $r$ = 0.86, respectively).<br><br>These results provide favorable evidence for the validity of the ASER assessment, the ORF assessment, and the reading comprehension assessment. |
| **Kannada (baseline sample)** | 1,815 | Not applicable | ASER (n=465) and ORF (n=461):<br><br>Inter-rater reliability for the ASER reading assessment is high (Cohen's kappa coefficient = 0.86). However, given the limited variation in reading levels and the sparse data in some categories, we cautiously conclude that the inter-rater reliability for the ASER reading assessment is good. | The Cronbach's alpha estimate for all items is 0.59 (n=514). This estimate improves substantially when comprehension question 1 is excluded (alpha =0.73). The inter-item correlation for question 1 is weak and it appears that it is a poorly performing item for this sample and the most difficult item with just 14% of students recording a correct response. All of this suggests that question 1 | The concurrent validity coefficient between the ASER and ORF assessment is in the high-moderate range ($r$ = 0.70, n=1815), and the validity coefficients of the reading comprehension assessment with the reading assessments are low ($r$ = 0.35 and 0.44, n=514). These coefficients are most likely attenuated given the presence of high floor |

| Language | n | Order of Test Administration[11] | Inter-rater Reliability | Reliability of the Reading Comprehension Component | Validity |
|---|---|---|---|---|---|
| | | | The inter-rater reliability between monitor and examiner for noting down words read correctly on the ORF assessment is high (Pearson's correlation coefficient = 0.96). The mean absolute error between the examiners' and monitors' reporting of words read correctly is less than a word (MAE = 0.2). This suggests a high degree of inter-rater reliability for the reporting of words read correctly. | should be excluded. However, we are cautious in interpreting the overall estimates for reliability as very few students were eligible to respond to three of the five comprehension questions, i.e., Q3 (n=20), Q4 (n=10), Q5 (n=7)). Despite the smaller sample size of the pilot study (n=372), more students responded to the comprehension questions due to higher levels of reading attainments. The pilot samples ranged from a high of 182 for the first question to 126 for the fifth question. Hence, the reliability estimates based on the pilot sample appear to be more robust. | effects. Average fluency rates increase from 0 for beginners to 36 cwpm for students categorized at the Standard 2-level text on the ASER assessment. The overall trends are in keeping with those observed in the pilot study, and together these findings provide favorable evidence for concurrent validity for the ASER and ORF assessments. |
| **Marathi (pilot sample)** | 366 | Results indicate that ASER and ORF are strongly correlated and the magnitudes of the correlation coefficients for Group A (n=184) are not significantly different from the coefficients of Group B (n=182).<br><br>We conclude there are no test order effects on student performance. | Reading Comprehension (Q1 n=235, Q2 n=200, Q3 n=148, Q4 n=79, Q5 n=41) and ORF (n=255):<br><br>The inter-rater reliability estimates for reading errors and the comprehension questions are almost perfect. The number of students administered the successive comprehension questions decreases, and the sample sizes for comprehension questions 4 and 5 are relatively smaller. However, these estimates suggest that there is high agreement (and low ambiguity) when recording errors or evaluating student responses to the comprehension questions, including the fifth inferential question. | The inter-item correlations between all the comprehension questions range from low-moderate (0.5) to strong (0.8). The overall Cronbach's alpha coefficient for the reading comprehension component is 0.71 (n=369). Typically, reliability coefficients of 0.70 and higher are considered adequate.<br><br>The reading comprehension questions appear to be performing well, and reliability based on internal-consistency is adequate. | The concurrent validity coefficients for ASER and ORF are moderately strong ($r$ = 0.74, n=366) and for the reading comprehension assessment with ASER is moderate ($r$ = 0.61, n=341) and with ORF is strong ($r$ = 0.87, n=341).<br><br>Average fluency rates increase from 0 for beginners to 49 cwpm for students categorized at the Standard 2-level text on the ASER assessment.<br><br>Overall, the results provide favorable evidence for the concurrent validity of the ASER assessment, the ORF assessment, and reading comprehension assessments. |

| Language | n | Order of Test Administration[11] | Inter-rater Reliability | Reliability of the Reading Comprehension Component | Validity |
|---|---|---|---|---|---|
| **Marathi (baseline sample)** | 1,464 | Not applicable | ASER (n=268) and ORF (n=267):<br><br>Inter-rater reliability for the ASER reading assessment is high (Cohen's kappa coefficient = 0.92). However, given the limited variation in reading levels and the sparse data in some categories, we cautiously conclude that the inter-rater reliability for the ASER reading assessment is good.<br><br>The inter-rater reliability between monitor and examiner for noting down words read correctly on the ORF assessment is high (Pearson's correlation coefficient = 0.93). The mean absolute error between the examiners' and monitors' reporting of words read correctly is about 2 words (MAE = 2.1). This suggests a high degree of inter-rater reliability for the reporting of words read correctly. | The Cronbach's alpha coefficient for the reading comprehension component is 0.68. The internal consistency estimate for the Marathi comprehension assessment is at the margin and very similar to the estimate reported for the pilot sample (Cronbach's coefficient alpha = 0.71). | The concurrent validity coefficients between the ASER and the ORF assessment are fairly high ($r$ = 0.74, p<.05, n=1463), while the associations of the reading assessments with the reading comprehension assessment are in the moderate to strong range ($r$ = 0.60 and 0.86, p<.05, n=1337). In addition, the average fluency rates progressively increase (from 3 to 43 cwpm) with higher reading levels on the ASER assessment. These high levels of mapping across the two reading assessments provide additional and favorable evidence for concurrent validity.<br><br>The findings from the baseline sample are similar to the trends noted for the pilot study sample. |
| **Oriya (pilot sample)** | 408 | Results indicate that ASER and ORF are strongly correlated and the magnitudes of the correlation coefficients for Group A (n=216) are not significantly different from the coefficients of Group B (n=192).<br><br>We conclude there are no test order effects on student performance. | Reading Comprehension (Q1 n=122, Q2 n=97, Q3 n=65, Q4 n=40, Q5 n=18) and ORF (n=196):<br><br>The inter-rater reliability estimates for reading errors and the comprehension questions are close to perfect. The number of students administered the successive comprehension questions decreases, and the sample sizes for comprehension | The inter-item correlations between all the comprehension questions range from low-moderate (0.5) to strong (0.7). The overall Cronbach's alpha coefficient for the reading comprehension component is 0.6 (n=240). Typically, reliability coefficients of 0.70 and higher are considered adequate.<br><br>Reliability based on internal-consistency for the reading | The concurrent validity coefficient for ASER and ORF is high ($r$ = 0.89, n=408) and for the reading comprehension assessment with ASER is moderate ($r$ = 0.51, n=247) and with ORF is moderately strong ($r$ = .74), respectively.<br><br>Average fluency rates increase from 0 for beginners to 51 cwpm for students categorized at the Standard 2-level text on the ASER assessment. |

| Language | n | Order of Test Administration[11] | Inter-rater Reliability | Reliability of the Reading Comprehension Component | Validity |
|---|---|---|---|---|---|
| | | | questions 4 and 5 are relatively smaller. However, these estimates suggest that there is high agreement (and low ambiguity) when recording errors or evaluating student responses to the comprehension questions, including the fifth inferential question. | comprehension assessment is below adequate. | Overall, the results provide favorable evidence for the concurrent validity of the ASER assessment, the ORF assessment, and reading comprehension assessments. |
| **Oriya (baseline sample)** | 901 | Not applicable | ASER (n= 263) and ORF (n= 263):<br><br>Inter-rater reliability for the ASER reading assessment is high (Cohen's kappa coefficient = 0.97). However, given the limited variation in reading levels and the sparse data in some categories, we cautiously conclude that the inter-rater reliability for the ASER reading assessment is good.<br><br>The inter-rater reliability between monitor and examiner for noting down words read correctly on the ORF assessment is high (Pearson's correlation coefficient = 0.99). The mean absolute error between the examiners' and monitors' reporting of words read correctly is less than a word (MAE = 0.1). This suggests a high degree of inter-rater reliability for the reporting of words read correctly. | The Cronbach's alpha coefficient for the reading comprehension component is 0.62 (n=391), which is similar to the Cronbach's alpha coefficient noted for the pilot sample (alpha = 0.60, n=240). Typically, reliability coefficients of 0.70 and higher are considered adequate. It is unclear whether the low estimates for internal consistency are attributable to the comprehension questions per se or to the low reading attainments of this group (average fluency rate = 19 cwpm). Reliability based on internal-consistency for the reading comprehension assessment is less than optimal, which is similar to the pilot study findings. | The concurrent validity coefficients between the ASER and the ORF assessment are high ($r$ = 0.85, p<.05, n=901), while the associations of the reading assessments with the reading comprehension assessment are in the moderate range ($r$ = 0.52 and 0.63, p<.05, n=514). These coefficients are most likely attenuated given the presence of a substantial number of non-text readers. However, the average fluency rates progressively increase (from 0 to 33 cwpm), with higher reading levels on the ASER assessment. These high levels of mapping across the two reading assessments provide additional and favorable evidence for concurrent validity.<br><br>These findings are similar to the findings of the pilot study. |

Full reports for each language provided upon request.

# Annex G: Data Collector Training and Data Collection Details

**Table G-1** provides details on the content covered during the 3-day master and state-level data collector trainings.

**Table G-1.    Master and state-level data collector training details**

| Day 1 | Day 2 | Day 3 |
|---|---|---|
| About USAID and its projects | Pilot field practice: practicing all the processes learnt on Day 1 | Quiz: Audio quiz to measure participant accuracy<br><br>Quiz: Revision and clarifications |
| Introduction to the evaluation | Feedback session | Monitoring and recheck |
| Assessment process:<br><br>• Preparation before going to school<br><br>• Collecting school information<br><br>• Sampling students<br><br>• Completing the child information and child consent forms and questionnaire<br><br>• Annual Status of Education Report (ASER) Testing<br><br>• Oral reading fluency (ORF) reading and comprehension testing<br><br>• Demonstration and discussion on the tool | Quiz: Assessment process | State-level training and assessment planning |

*School-based practice.* 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.

*Data collector performance.* Central- and state-level trainings included an Assessor Accuracy Measure (AAM) to evaluate participants' proficiency with the data collection processes and to ensure that different data collectors had the ability to score the same child accurately on the ORF subtask, thus increasing the rigor and reliability of the measurement. Participants listened to an audio recording of a child reading the passage and answering the comprehension questions. Data from the AAM 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 through a paper-based quiz.

All master trainers scored above 85% on both the Annual Status of Education Report (ASER) quiz and early grade reading assessment (EGRA) AAM.

Participants from the state-level trainings scoring less than 70% on the assessment process 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 G-1** presents master trainers scores on the ASER quiz and EGRA AAM.

**Figure G-1. Master trainer scores on ASER quiz and EGRA AAM**



Participants – Assessment Process Quiz Score

Participants – ORF Audio Quiz Score
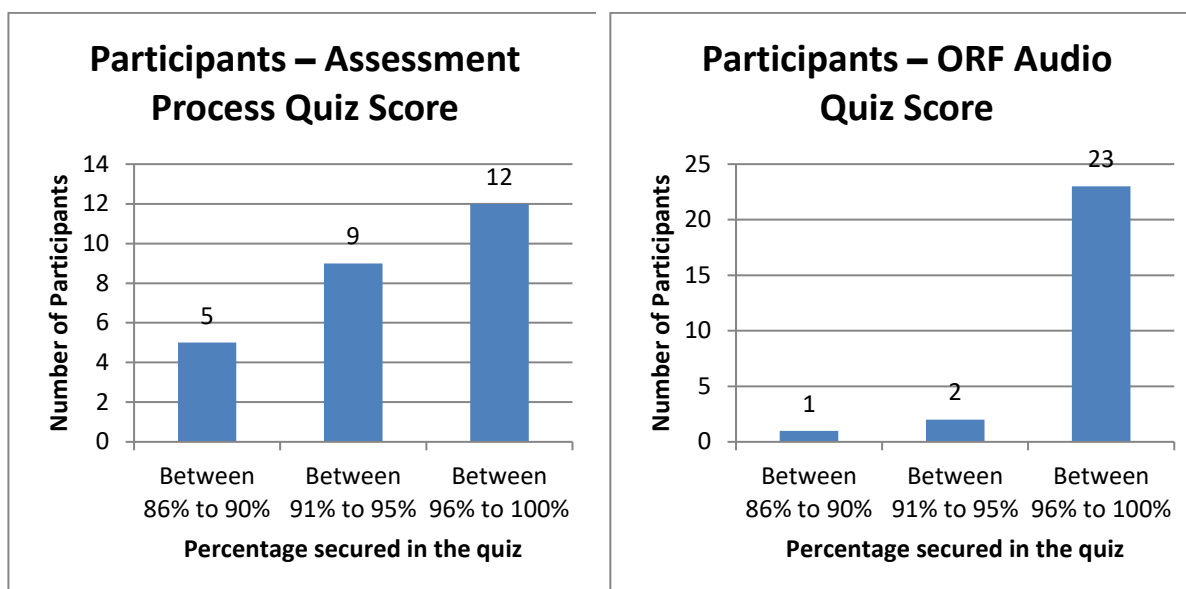
*Figures G-2* and *G-3* present data collectors' scores by location on the ASER quiz and the EGRA AAM.

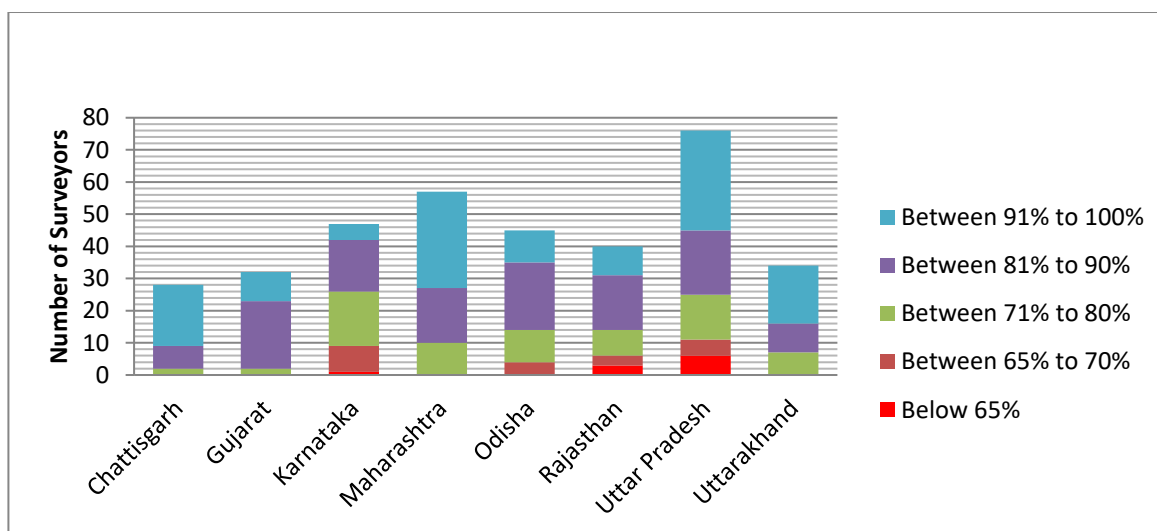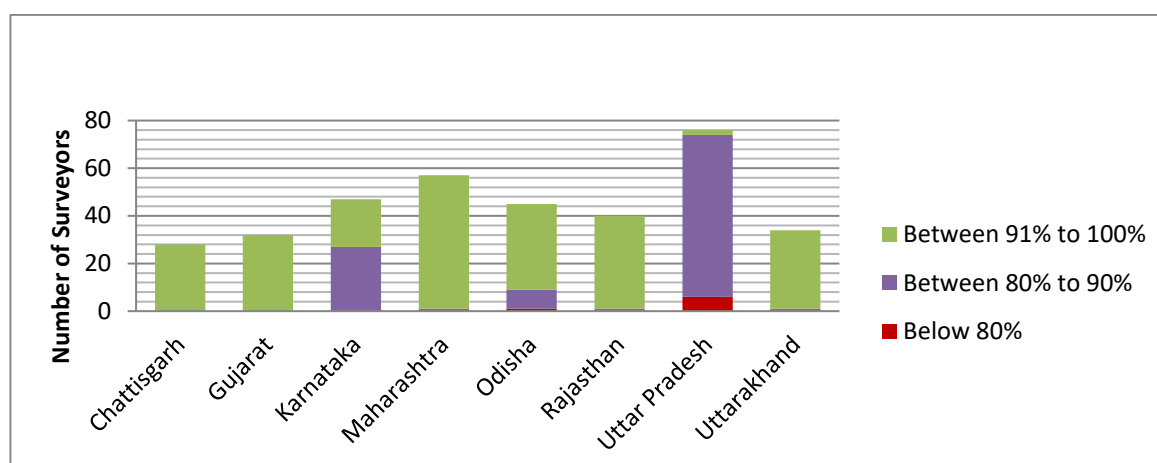**Figure G-2. Data collector assessment process quiz scores**

**Figure G-3. Data collector ORF audio quiz scores**



**Table G-2** presents the number of data collectors and monitors selected by state.

**Table G-2. Number of data collectors and monitors selected, by state**

| State | Participants on Day 1 | Selected as Data Collectors | Selected as Monitors | Not Selected |
|---|---|---|---|---|
| Chhattisgarh | 30 | 22 | 6 | 2 |
| Gujarat | 32 | 26 | 6 | 0 |
| Karnataka | 52 | 47 | 10 | 5 |
| Maharashtra | 59 | 46 | 11 | 2 |
| Odisha | 53 | 33 | 12 | 8 |
| Rajasthan | 49 | 32 | 8 | 9 |
| Uttar Pradesh | 84 | 58 | 18 | 8 |
| Uttarakhand | 38 | 25 | 9 | 4 |
| **TOTAL** | **397** | **289** | **80** | **38** |

Each field team was made up of the following people: Data Collector, Monitor, and Supervisor/ASER State Team. Based on participant performance during training, about 25% of the participants in every state were given the role of "Monitor." Monitors visited a school with a different data collector each day and aided in ensuring data quality. Supervisors included the ASER state teams along with Pratham's local staff (master trainers) in respective states/districts, individuals belonging to external organizations, and those who have helped in execution of ASER surveys in the past. Supervisors were responsible for the successful execution of assessment in the states, including all quality control processes—monitoring and conducting desk and field checks.

At a minimum, 120 schools (60 treatment and 60 control) were visited per project location; as necessary, additional schools were sampled and visited in some states where enrollment and attendance were low. In each school, assessments were carried out by one data collector. Upon reaching the school, the data collector met the head master and secured permission to carry out the assessment in the school. Data collectors were asked to carry out simple activities with students in the class, such as introduce themselves to the students

or play a quick game before sampling or beginning the assessment. This was done to ensure that students felt at ease with the data collectors during the assessment. Next, data collectors obtained information on enrollment and attendance for Standard 2. Then, 20 students (10 boys and 10 girls) were randomly sampled for the assessment. Each assessment was conducted outside of the classroom to avoid any distraction from other students or the teacher.

Data collection timelines for each project location varied depending on getting permission to carry out the survey in schools as well as local holidays due to festivals, etc. *Table G-3* shows dates for data collector training and fieldwork by each project location.

**Table G-3.    Training and data collection timelines by project location**

| Name of the Program | Training Dates | Data Collection Dates |
|---|---|---|
| Scaling Up Early Learning Intervention – Uttarakhand | September 13–15 | September 16–October 13 |
| Scaling Up Early Learning Intervention – Chhattisgarh | September 14–16 | September 18–25 |
| Nurturing Early Literacy – Rajasthan | September 13–15 | September 16–25 |
| Teacher Innovations in Practice – Uttar Pradesh | September 13–15 | September 16–October 7 |
| Start Early: Read in Time – Uttar Pradesh | September 13–15 | September 16–October 9 |
| Start Early: Read in Time–Odisha | September 15–17 | September 17–25 |
| Nurturing Early Literacy – Karnataka | September 14–17 | September 19–23 |
| Nurturing Early Literacy – Maharashtra | September 13–15 | September 16–October 14 |
| Right to Read – Maharashtra | September 13–15 | September 16–October 14 |
| School Excellence Program –- Gujarat[12] | September 11–13 | Not collected |
| Right to Read – West Bengal[13] | January 2018 | Not collected |

[12] School Excellence Program – Gujarat: State-level training for this project location was successfully completed from September 11–13, 2017, in Surat, Gujarat; however, data collection was not conducted for this project due to lack of permission to access schools from the local government.

[13] Data collection was not conducted for this project due to lack of permission to access schools from the local government.

# Annex H: Number of Schools Monitored and Checked

Field teams were comprised of data collectors, monitors, and supervisors/ASER state teams. About 25% of training participants in every state were given the role of monitor based on their performance during the training. Monitors rotated between schools and accompanied a different data collector to ensure data quality. In addition to this, a third layer of quality control was created through the supervisor role. Supervisors were responsible for the successful execution of assessments in each state, including all quality control processes – monitoring, desk and field checks. Supervisors included the ASER state teams along with Pratham's master trainers in their respective states/districts, as well as individuals belonging to external organizations who have been involved with past ASER surveys.

**Table H-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 1 school. This includes talking to the Head Master and explaining the evaluation, collecting school information, building rapport with Standard 2 students in a big group, sampling students for assessment and completing the assessment with the sampled students.<br>Each data collector is assigned 3 to 6 schools based on the size of the team. | Visits 1 school per day with 1 data collector and fills in the data collector booklet (marked as "monitor" in type) for the visited school. 20–30% schools were monitored in this way. | Coordinates with monitors and data collectors by phone.<br>Conducts monitoring visits to schools during data collection and fills out a monitoring format during field visits. This monitoring 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. | Desk and field checks.<br>Once the hardcopies are submitted they are thoroughly checked for completeness and correctness. Based on the feedback of the desk check, schools are picked for field check. During field check, the supervisors/ASER state team members speak to Head Masters and students and ask questions to assure data collection was done following the correct protocol. |

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

**Table H-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 Reading Intervention – Uttarakhand | 90 | 90 | 29 | 28 | 21 | 11 | 90 | 90 | 12 | 23 |
| Scaling Up Early Reading Intervention – Chhattisgarh | 60 | 60 | 15 | 19 | 17 | 21 | 60 | 60 | 21 | 10 |
| Teacher Innovations in Practice – Uttar Pradesh | 70 | 70 | 21 | 26 | 16 | 9 | 70 | 70 | 7 | 5 |
| Start Early: Read in Time – Uttar Pradesh | 70 | 70 | 33 | 20 | 19 | 21 | 70 | 70 | 7 | 8 |
| Start Early: Read in Time – Odisha | 60 | 60 | 14 | 21 | 3 | 2 | 60 | 60 | 3 | 8 |
| Nurturing Early Literacy – Rajasthan | 60 | 60 | 18 | 12 | 17 | 15 | 60 | 60 | 14 | 16 |
| Nurturing Early Literacy – Karnataka | 60 | 60 | 13 | 17 | 3 | 2 | 60 | 60 | 15 | 10 |
| Nurturing Early Literacy – Maharashtra | 70 | 70 | 19 | 17 | 2 | 2 | 70 | 70 | 20 | 7 |
| Right to Read – Maharashtra | 67 | 44 | 25 | 19 | 17 | 7 | 69 | 44 | 13 | 7 |
| TOTAL | 607 | 584 | 187 | 179 | 115 | 90 | 609 | 584 | 112 | 94 |

# Annex I: Comparing Treatment and Control Schools

| VARIABLES | Scaling Up Early Reading Intervention - Uttarakhand | Scaling Up Early Reading Intervention - Chhattisgarh | Teacher Innovations in Practice - Uttar Pradesh | Start Early: Read in Time - Uttar Pradesh | Start Early: Read in Time - Odisha | Nurturing Early Literacy - Rajasthan | Nurturing Early Literacy - Karnataka | Nurturing Early Literacy - Maharashtra | Right to Read - Maharashtra |
|---|---|---|---|---|---|---|---|---|---|
| Control | -9.187*** | -10.27*** | -1.259 | -0.350 | -7.909** | 0.880 | -0.204 | 3.182 | -0.109 |
|  | (3.118) | (2.935) | (1.806) | (1.407) | (3.739) | (0.629) | (0.744) | (5.456) | (0) |
| Constant | 20.21*** | 17.04*** | 6.137*** | 3.889*** | 22.29*** | 0.329 | 1.572*** | 53.94*** | 0.109 |
|  | (2.391) | (2.551) | (1.138) | (0.714) | (2.651) | (0.234) | (0.504) | (3.614) | (0) |
|  |  |  |  |  |  |  |  |  |  |
| Observations | 1,681 | 1,882 | 1,763 | 1,771 | 903 | 1,257 | 1,821 | 1,469 | 1815 |
| R-squared | 0.016 | 0.025 | 0.001 | 0.000 | 0.009 | 0.002 | 0.000 | 0.001 | 0.001 |

Standard errors in parentheses

*** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level

Treatment significantly different from Control

| All | Yes + | Yes + | No | No | Yes + | No | No | No | No |
|---|---|---|---|---|---|---|---|---|---|

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

| All | Yes + | Yes + |  |  | No |  |  |  |  |
|---|---|---|---|---|---|---|---|---|---|