

Improving Early Grade Reading Performance: Results of School Based and Community-based Interventions in Bangladesh

Ranak Chandra Mohanta
Shahana Parvin Lata
Emily Richardson
Shahin Islam
Md. Akidul Islam

Abstract

Reading Enhancement for Advancing Development (READ), which started in October 2013, is a 5-year project of Government of Bangladesh, supported by the US Agency for International Development (USAID) and implemented by Save the Children International (SCI). READ aims to improve early grade Bangla reading competencies among Grade I-III students. READ Core model provides three school-based interventions. The READ Core Plus model adds a fourth intervention: Community Reading Camps. The randomized controlled trial (RCT) study has employed a longitudinal approach by conducting baseline (May'15), midline (April'16) and endline (Jul'17) assessments.

Findings show there is a large improvement in all the literacy sub-skills in both READ Core and READ Core Plus groups. For emergent and decoding skills, students performed 2.5 times better in the endline compared to the baseline. Importantly, above 90% students from both school-based and community-based intervention groups could correctly read more than 60 words per minute. On average, students answered more than 80% of the comprehension questions correctly.

Interestingly, there was no significant additional gain in achievement from community-based reading camp intervention. Girls performed significantly better than boys in many lower and higher orders literacy skills. Students with lower socio-economic background and those with fewer books at home performed poorly compared to those from higher socio-economic background and with more books at home. The interventions were not a sufficient compensation for socio-economic disadvantage of students. There was also no difference in results based on students' earlier attendance in preschool, perhaps a commentary on the quality of the preschool activities.

Key Words: Community-based reading intervention, School-based reading intervention, Foundational skills in primary education, Primary education in Bangladesh

1. Introduction

Bangladesh has made dramatic strides in improving access to basic education, driven by strong government leadership and successful partnerships among government, donors and NGOs. However, the government's National Student Assessment 2011 points to weak Bangla competencies among the grade III and V students. The findings also indicate that Bangla competencies for grade level decreases from grade III to grade V, suggesting that children fall more and more behind due to a weak foundation. This weak Bangla competency has a detrimental effect on the entire education system and the economy as a whole.

The Government of Bangladesh's Third Primary Education Development Program (PEDP III) is aimed at addressing the quality challenges. The Reading Enhancement for Advancing Development (READ) project, implemented by Save the Children International (SCI), aims to improve early grade Bangla reading competencies among Grade I-III students. READ has a target to directly benefit an estimated 1.1 million Grade I-III students in 5,112 READ-supported schools in selected districts of Bangladesh. READ expects to reach an estimated 4 million direct and indirect beneficiaries in seven divisions of the country. Indirect beneficiaries include grade IV and V students in target schools, people reached through public awareness activities on reading, and teachers as well as students from non-READ schools who would benefit from resources on the READ website. Between 2014 and 2017, READ project directly served 5,112 schools in 51 upazilas of 19 districts across 7 divisions.

To measure the value added of the READ Core Plus model on children (the community aspect) compared to the school-based READ Core model, READ team, in collaboration with Silvia Diazgranados Ferráns, a PhD candidate at Harvard University, designed and implemented a study using a multi-site randomized controlled clustered evaluation. Baseline and midline data were collected in 2015 and 2016 respectively. Innovations for Poverty Action (IPA), as an external evaluator, has conducted the endline reading assessment for Grade III students in July 2017. The primary objectives of the evaluation were:

- a. Administering surveys using structured questionnaires to collect data at the student, household and school levels.
- b. Evaluating the additional impact of READ Core Plus compared to READ Core on student outcomes through rigorous empirical analysis of the data collected. In particular, the study analyzed the impact of READ Core Plus on Bangla reading skills compared to READ Core students and what factors (income, gender, region, etc.) might have caused impact heterogeneity or differences in project outcomes.

Following activities were undertaken to implement the evaluation:

1. Planning: In collaboration with SCI, field-administration plan for endline data collection was developed to ensure the quality of the data collected. The plan specified the timeline, training, logistics, data entry and supervision. IPA developed Stata package - 'High Frequency Check' was run daily to ensure reliability of the incoming data.
2. Training: In close collaboration with SCI, supervisors and enumerators were trained for 4 days including a day of field-practice in 3 schools in Manikganj district. Education

Specialist from SCI also introduced the enumerators to the right pronunciation of the letters and words in the assessments. Besides repeated mock-interviews, enumerators also received feedbacks after the field-practice to ensure clear understanding of the data-collection protocol for the endline.

3. Data collection: According to the plan, the endline survey was administered between 10th July and 18th July 2017. The field-work was followed by data cleaning in preparation of a clean data set for analysis.
4. Analysis: Under the direct supervision of the Principal Investigator, Professor Atonu Rabbani from Dhaka University and in close collaboration with SCI, an analysis plan was prepared and followed to prepare the evaluation report. The report was revised based on discussions between stakeholders.
5. Reporting: Final report was submitted to SCI on 23rd November 2017.

This article summarizes the evaluation report which is organized in eight parts including introduction and background, research methodology, analysis, results and findings, discussions and conclusions.

2. Background

Despite some remarkable achievements, Bangladesh is seriously lagging in maintaining the quality of education. One of such weak areas is early grade reading competence. A national survey done by the government in 2011¹ identified that a third of the Grade III students did not have the relevant level of Bangla reading competence. A subsequent survey conducted in 2013² revealed that there was very little improvement in reading competence since the earlier survey.

Early grade reading competence is critical to the overall educational outcome of a child. Researchers have found that early grade students who fall behind in reading skills compared to their peers perform progressively worse in later Grades.³ Another study among children from low-income families found that the first grade students who differed in reading skill diverged further by the time they reached fourth Grade and whether a child would improve was established very early -- by second Grade.⁴ First Grade reading competence can have a strong impact on the overall literacy of a child (e.g. vocabulary,

¹ 2011 National Student Assessment for Grades 3 and 5 National Report, Government of the People's Republic of Bangladesh Directorate of Primary Education, 18 December 2014

² 2013 National Student Assessment for Grades 3 and 5 National Report, Government of the People's Republic of Bangladesh Directorate of Primary Education, 29 December 2012

³ Crouch, L. 2012. Why Early Grade Reading: An Economist's Perspective. Presentation given at 'All Children Reading Workshop', Kigali, Rwanda, 28 February 2012

⁴ Predicting Improvement After First-Grade Reading Difficulties: The Effects of Oral Language, Emergent Literacy, and Behavior Skills. Spira, Elana Greenfield; Bracken, Stacey Storch; Fischel, Janet E. *Developmental Psychology*, Vol 41(1), Jan 2005, 225-234. <http://dx.doi.org/10.1037/0012-1649.41.1.225>

reading comprehension, general knowledge) even after ten years,⁵ which means that the impact can last a lifetime. A study among 4,000 students in the United States found that students who cannot read proficiently by the end of third grade are four times more likely to drop out without a high school diploma.⁶ Early grade poor readers are also more likely to repeat grades.⁷ As such, poor reading competence among early grade students in Bangladesh is a matter of grave concern.

In this context, Save the Children started a five-year program called Reading Enhancement for Advancing Development (READ), funded by the US Agency for International Development (USAID). READ incorporates evidence-based practices that emphasize capacity building and sustainability, focusing on four areas of intervention: 1) teacher education and continuous professional development; 2) reading assessment; 3) increased availability of reading material, and 4) increased opportunities in the community to read and to provide support to beginning readers by those outside the school. Specifically, the READ model provides participating schools with evidence-based strategies to assess and teach reading skills such as letter knowledge, phonemic awareness, vocabulary, fluency and comprehension. 'READ Core package' was offered at the early stage of the project, addressing school input components (first three intervention areas). Read Core package includes:

- Teacher training on Early Grade Reading Instruction
- Teacher training on Early Grade Classroom Reading Diagnostic and Formative Assessment Training of Administrators (Head Teachers, Upazila Education Officers (UEO), Upazila Resource Centre (URC) Instructors and Assistant Upazila Education Officers (AUEO)) on academic supervision.
- School-based book bank: READ supplied 70 titles and one book stand for all schools. The book stands are designed to display books and bring them within easy reach of children. The books are visible through transparent pockets so that children will be attracted to them, according to their diverse interests. Books are color-coded according to reading level so that children can choose books according to their reading skill. Each class has a specific time to receive and read books from the book corner.
- Orientation of School Management Committees to the project and where they can help, such as using local resources (School Level Improvement Plan) to acquire print material/books.

⁵ Early reading acquisition and its relation to reading experience and ability 10 years later. Cunningham, Anne E.; Stanovich, Keith E. *Developmental Psychology*, Vol 33(6), Nov 1997, 934-945. <http://dx.doi.org/10.1037/0012-1649.33.6.934>

⁶ Hernandez, D. 2011. *Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation*. New York, USA: The Annie E. Casey Foundation

⁷ Annie E. Casey Foundation. 2010. *Early Warning: Why Reading By the End of Third Grade Matters*. Baltimore, MD, USA: Annie E. Casey Foundation.

In 'READ Core Plus', the community component is added on top of the Read Core package. Students in Core Plus schools benefit from all the above-mentioned activities as well as an access to a Community Reading Camp. Reading camps provide children with an opportunity to enhance their reading competence by participating in fun, engaging in literacy activities outside the school environment. Two reading camps are established within each school catchment area, comprising of both Grade I and II students. There are 60 students in each camp. Camps are held once a week for 8 months in a year (30-32 sessions). Each session is 90-minute long in which two Community Literacy Volunteers (CLVs), who receive a two-day training and a small stipend, engage students in activities such as games, lessons and comprehension-telling through which they can practice their literacy skills. Parents and community members in each READ community program are actively involved in managing the program, supporting the CLVs, and maintaining materials.

The following is a description of activities in the camp:

- **Community-based book bank:** To ensure that the children have access to adequate and diverse reading materials appropriate to their level of reading outside school, each reading camp is provided with a collection of 70 illustrated books, 2 alphabet primers, 2 language games, and guidelines on how to read for children and how to read with children. The children can borrow one book per week. CLVs oversee the book bank.
- **Reading Buddies:** To provide need-based support to beginner or younger readers and to develop a habit of reading for pleasure among all participating students, CLVs match pairs of students who live close by but have different levels of literacy skills. Older buddies or those with high levels of fluency and reading comprehension are trained on how to read to the younger readers. The younger reader is encouraged to borrow books from the book bank and read them together with his or her 'Buddy' who is more advanced. Books are distributed at the reading camp, but the actual reading takes place outside the camp, at home or elsewhere in the community.
- **Comprehension telling by community resource person:** A community resource person is hired to tell stories to the children. A comprehension telling session is held twice a month. Comprehension tellers may include parents who attend the parenting session, grandparents or anyone else from the community. Stories are selected from oral traditions, cultures, books, poems, comprehension or may be made up and are well-suited to the social context.
- **Reading to children:** In each session, the CLVs read for comprehension from the book bank to the children. The CLV asks questions during and after the reading.
- **Parent Awareness Sessions:** Parents are invited to the camp once a month for seven months to introduce them to concrete activities to improve the language development and literacy of their children at home. Separate 60-90 minute sessions are organized for about 20-30 participating parents.

3. Research Methodology

Study Design. The study uses a multi-site randomized controlled clustered evaluation to measure the value added of the READ Core Plus model for children (the community aspect) compared to the school-based READ Core model. Specifically, the study aimed to answer the following research questions:

- Does receiving an offer to attend READ Core Plus improve students' literacy skills as compared to students who only received an offer to attend READ Core?
- Does the effect of the offer to receive READ Core Plus vary for students according to their gender and the division in which their school is located?

School Selection. To select the sample of our study, the research team first randomly chose 70 READ Core Plus Government Primary schools (GPS) from among the 80 that were selected to receive the intervention. For each treatment school, the researchers then located the closest GPS receiving READ Core within the same Upazila, to be used as a control group.

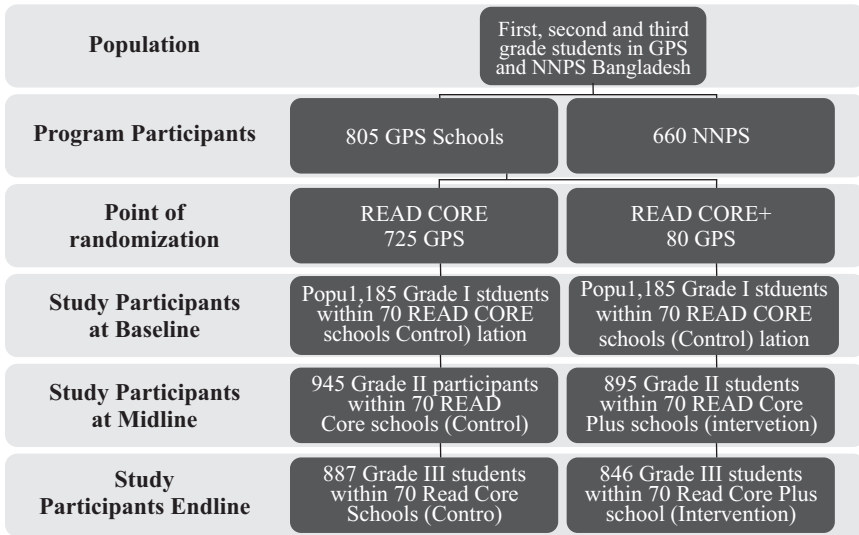
Student Selection. Within each selected site, the researchers randomly selected 20 first grade students to participate in the study. Given that some schools had less than 20 first grade students, the final study sample was comprised of 2,317 first grade students in 140 GPS schools: 1,185 students in 70 schools receiving the READ Core Package, and 1,134 students in 70 schools receiving the READ Core plus Package. At midline, researchers tracked down all the participants and were able to find 1,840 students.

Because the attrition was high during the midline, extra effort was made by SCI team to ensure student attendance during endline survey to track the children. Finally, we could find 1,733 students in the endline, which represent an attrition rate of 25.2%. The potential effects of the attrition is discussed in 'balance and attrition analysis' section. With replacement, we interviewed 2,027 students (male: 932, female: 1,095) in total. This report presents results only for students who were present in both baseline and endline, including repeat students. We include students who repeated Grade II last year as it is unlikely that grade repeaters would be excluded from project implementation in practice.

Data Collection. Data were collected in three rounds, baseline, midline and endline, respectively in May-June, 2015, March-April, 2016, and August-September, 2017.

The study design and sampling plan are shown in Figure 1.

Figure 1: Study Design and Sampling Plan



Baseline and end-line samples and their characteristics are shown in Table 1.

Table 1: Baseline and Endline Study Sample by Treatment Status, Sex and Division

	Baseline					Matched Endline				
	READ Core		READ Core Plus		Total	READ Core		READ Core Plus		Total
	Male	Female	Male	Female		Male	Female	Male	Female	
Barisal	106	97	89	96	388	88	82	61	83	314
Chittagong	40	51	36	37	164	28	41	21	28	118
Dhaka	90	81	116	130	417	50	59	67	84	260
Khulna	150	141	104	112	507	117	116	80	93	406
Rangpur	122	131	105	91	449	83	104	82	88	357
Sylhet	89	87	112	106	394	53	66	79	80	278
Total	597	588	562	572	2319	419	468	390	456	1733

4. Quality Control

We used Cronbach’s alpha test to measure internal consistency of the assessment. We found the overall Cronbach’s alpha to be 0.92, indicating high level internal consistency of the tests. It means that each component of the assessment test measures ‘early grade reading’ skills correctly and reliably. We also found the assessment test for Grade III used in this test to be unidimensional (only one factor has eigen value of more than 1), which means that the test is

testing a single overall construct - “early grade reading”. To ensure data validity, we used ‘High Frequency Check’ every day to ensure that incoming data meets the required criteria such as: non-empty value for unescapable variables; logic; age within a low and high limit, etc.

About ten percent children from the overall sample were interviewed for inter-rater reliability (IRR). This sample size for IRR should be enough as prescribed by EGRA toolkit.⁸ A single student was interviewed by two enumerators at a similar time, where one enumerator remained quiet. We did not detect any significant differences (no p-value is smaller than 0.1) between the data collected by two enumerators using a t-test to measure inter-rater reliability or how reliable the collected data is.

The enumerators and supervisors were selected from the experienced pool of IPA surveyors and most of them have prior experience in administering the EGRA reading assessment using the Tangerine software. Each supervisor, an experienced enumerator, supervised a team of six to seven enumerators.

In close collaboration with SCI, supervisors and enumerators were trained for 4 days including a day of field-practice in 3 schools in Manikganj district. Education Specialist from SCI also introduced the enumerators to the correct pronunciation of the letters and words in the assessments. Beside repeated mock-interviews, enumerators also received feedbacks after the field-practice to ensure clear understanding of the data-collection protocol for the baseline.

5. Analysis

5.1 Balance and Attrition Analysis

From the midline report, we saw that there were no statistically significant differences between the Read Core and Read Core Plus students on a range of exogenous⁹ and endogenous¹⁰ variables during baseline. Based on the balance analysis, it can be concluded that the groups were equivalent and had no differences in the observable characteristics that might affect the students’ performance before intervention. Therefore, any differences found in students’ performance can be attributed to program impact.

The sample size was 2,319 first grade students in 140 GPS schools in the baseline. We could track 1,733 students in Grade III in total. Our attrition rate from original sample is 25.2%, which is about eight percent each year. Attrition might change the existing balance in baseline between the two groups. If, balance of observable characters in baseline changes due to attrition, difference in students’ performance in endline cannot be solely attributed as impact of the program. Differences might be due to changes in those observable characters as well. We conducted a T-test to check whether the balance of the baseline characteristics of the students has changed due to attrition (Table 2).

⁸ A sample size of less than 100 for IRR will likely not yield useful information.

⁹ Exogeneous variables are factors that cannot be changed by the actions of the participant, such as: age, gender, school size etc.

¹⁰ Endogeneous variables are factors that can be changed by the actions of the participant, such as: test score, after-school activities etc.

Table 2: Balances in the Baseline Characteristics of the Treatment and Control Students that were Re-contacted and Lost at Endline

Dimension	Variable	Endline re-contacted students				Endline 'lost' students			
		Mean Core	Mean Core Plus	P-value from T-test	Number	Mean Core	Mean Core Plus	P-value from T-test	Number
Demographic information	Male	0.530	0.547	0.439	1677	0.416	0.409	0.844	642
	Age	6.644	6.748	0.185	1475	6.842	6.895	0.680	548
	SES	5.058	4.946	0.493	1676	4.894	4.545	0.086	641
	Home Literacy Materials	0.682	0.637	0.298	1732	0.634	0.599	0.531	697
Student Academic Trajectory	Attended ECD	0.510	0.537	0.558	1667	0.511	0.479	0.546	638
	Repeat Preschool	0.253	0.278	0.419	1660	0.251	0.244	0.879	631
	Repeat First	0.259	0.280	0.522	1661	0.370	0.282	0.113	633
	Missed school	0.415	0.425	0.795	1665	0.415	0.458	0.417	633
	Study time (Hours)	2.042	2.019	0.701	1487	1.938	2.002	0.355	575
Reading materials	At school	0.121	0.092	0.270	1732	0.109	0.088	0.475	697
	At community	0.088	0.114	0.273	1732	0.109	0.127	0.561	697
After school Activities	Doing homework	5.281	5.172	0.452	1611	4.933	5.217	0.166	613
	Attending Afterschool program	0.796	0.785	0.938	1070	0.806	0.503	0.149	407
	Private Tutoring	3.612	3.576	0.883	1569	3.307	3.337	0.922	597
	Reading books at home	5.802	5.726	0.675	1652	5.669	5.826	0.453	627
	Music, dancing, art	1.706	1.721	0.915	1532	1.430	1.436	0.981	580
	Organized sports	2.177	2.292	0.534	1565	2.544	2.362	0.467	599
	Religious activities	2.075	1.809	0.329	1543	2.045	2.515	0.161	584
	Watching TV	3.293	3.129	0.497	1609	3.233	2.895	0.267	614
	Hanging out with friends	3.737	3.650	0.642	1619	3.433	3.365	0.787	615
	Working at a job	0.504	0.310	0.025	1445	0.445	0.387	0.641	543
	Doing house chores	4.889	4.865	0.871	1649	4.633	4.792	0.487	626
	Taking care of siblings	3.264	3.096	0.355	1618	3.136	3.252	0.703	611

From Table 2, we can see that the lost students appear to differ from the tracked students only in case of two variables: socio-economic status and after school job, though students were balanced in these two characteristics in baseline. Among the lost students, Read Core students reported better socio-economic status compared to the Read Core plus students ($p < 0.1$), whereas there is no such significant difference between Core and Core plus among the tracked students. Among the tracked students, Read Core students reported having significantly more after-school jobs compared to the Read Core plus students ($p < 0.05$), whereas there is no such significant difference among the lost students. Overall, we see that there is no sign of systematic attrition that could potentially affect the outcome. Although, we might overestimate the causal effect of READ Core Plus activities as lost READ Core Plus students were more disadvantaged than the lost READ Core students, and as fewer re-contacted READ Core Plus students have after-school job than the re-contacted READ Core students.

5.2. Impact

In the previous section, we have seen that, there is no major statistically significant difference between the two groups in baseline. We have also seen that the re-contracted READ Core and READ Core Plus students in the endline are balanced in almost all the dimensions. As a result, we can safely assume that the treatment and control groups are balanced. Thus, we are applying the following random school intercept model, similar to the method used at midline, to determine the additional impact of READ Core Plus on the students' reading skills:

$$Y_i = \alpha + \beta X_i + \gamma S_i + \epsilon_i$$

where, Y_i is the outcome variable (reading tests: letter identification, similar beginning sounds etc.) of i th student, X_i is the treatment status of i th student, S_i is a vector of covariates (SES and gender) and ϵ_i is error terms.

Using the coefficient of multi-level regression analysis, we checked the interaction between treatment and students' socio-economic status (asset index), gender, reading materials at home, and attendance in ECD programs, to see if the treatment has similar impact across groups.

There are two limitations of this assessment model that should be noted.

1. This control group may not be representative of the overall sample population, as they were selected based on the nearby treatment schools, rather than randomly choosing control schools from the available sample. Actually, the possibility of these being similar in observable and unobservable characteristics to treatment schools was high.
2. EGRA test is administered following the established manual by SCI – Bangladesh. There are two main variation of standard EGRA recommendation in this test:
 - Timed task: In the current test format, only reading comprehension is a timed subtask. This enables us to measure fluency, but only for the Readers. As, fluency is categorized as one of the important skills for comprehending, decoding subtasks could also be timed for measurement. This should help us understand the relationship between fluency in decoding letters and ability to decode words/non-words for non-readers, as well as, students with intermediate comprehension skill.

- To be a successful reader, children must combine both decoding and sight-recognition skills; tests that do not include a decoding exercise can overestimate children's ability to read unfamiliar words (as the words tested may be a part of the sight-recognition vocabulary). (Gove 2011) In the current design, non-word subtask is not included in the test, thus we could not test their ability to read unfamiliar words.

6. Results

6.1 Reading Assessment¹¹

All learners of alphabet-based languages acquire reading skills in phases. Initially, students build basic literacy skills by developing letter-sound knowledge, word knowledge and simple decoding of letters into sounds. Gradually, students learn simple decoding of letters into sounds and they can identify and sort familiar words by creating a bank of words that students can recognize on sight. Gradually, students learn to read fluently and build a framework that enables them to attribute meaning to written text. (Gove 2011).

The EGRA tool has been designed based on the principle of acquiring reading skills in phases. EGRA measures the acquisition of reading skills in three progressively advanced phases: 1) Emergent literacy phase when a learner develops a phonemic awareness, 2) Decoding phase when a learner learns to identify letters, name syllables and read words and non-words, and 3) Confirmation and fluency phase when a learner develops oral reading fluency with comprehension.

The reading assessment tool used in this study includes the relevant subtasks associated with the above-mentioned phases as follows:

- 1st phase (emergent literacy skills): phonemic awareness (similar beginning sounds and similar ending sounds), and listening comprehension.
- 2nd phase (decoding): letter identification and familiar (most used) words reading.
- 3rd phase (fluency and comprehension¹²): reading comprehension (fluency, accuracy), and comprehension questions.

The result¹³ of the endline assessment is presented below as per these defined phases.

To identify whether there are significant differences in literacy outcomes of students who attended the READ Core Plus and those who only received READ Core, we conduct regression analysis, controlling for demographic characteristics—gender and asset index¹⁴. From this regression, we can see that there is no significant difference in performance between the Read Core and Read Core Plus students in any literary outcome (Table 3).

¹¹ Among the tracked students, 205 students repeated grade in the previous year. That means, these students are in Grade II during survey.

¹² EGRA Toolkit mentions this phase as 'confirmation and fluency'. To make it simpler for reader, we call this phase as 'fluency and comprehension'.

¹³ Results for reading skills were presented only for the students who were tracked, including students who repeated last year.

¹⁴ Asset index was created adding all the available asset options: electricity, refrigerator, television, buffalo/cow, goat/sheep, chicken/duck, land, bicycle, motorcycle, computer, mobile phone, and internet.

Table 3: Difference in Literacy Outcomes between READ Core and READ Core Plus Students Controlling for Gender and Asset Index

Literacy Outcome	Constant	Coefficient	S.Err	P-value	N
Emergent Literacy skills					
% similar beginning words correct	0.028	-0.006	0.015	0.674	1732
% ending rhyme in words correct	0.039	-0.013	0.017	0.426	1732
% listening comprehension answered correctly by non-readers (Story 1)	-0.055	-0.050	0.047	0.293	154
% listening comprehension answered correctly by non-readers (Story 2)	-0.049	0.022	0.065	0.736	122
Decoding					
% letter correct	0.011	-0.006	0.008	0.419	1732
% frequent/most used words correct	0.026	-0.012	0.018	0.512	1732
% antonym correct	0.012	-0.018	0.017	0.296	1732
Fluency and Comprehension					
% of students who are self-reliant reader (Story 1)	0.023	-0.015	0.021	0.473	1732
Accuracy score with missing for non-readers (Story 1)	0.013	0.007	0.006	0.217	1578
Fluency (Story 1)	6.388	1.006	2.529	0.689	1578
% of comprehension questions, correctly by readers (Story 1)	0.006	0.005	0.013	0.7	1578
% who answered more than 8 comprehensions correctly (all students) (Story 1)	0.029	-0.028	0.034	0.411	1732
% of students who are self-reliant reader (Story 2)	0.036	-0.020	0.019	0.276	1529

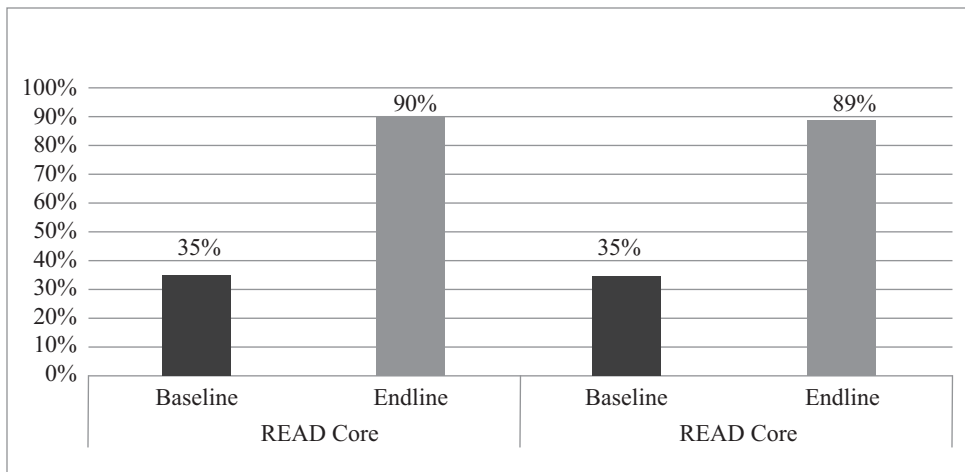
In the following section, we delineate the difference in means between Read Core and Read Core Plus students by each literacy outcome and discuss whether the differences are statistically significant. We also compare the endline outcomes to those in the baseline and observe a large improvement in performance from baseline in almost every subtask in both Core and Core Plus group.

6.2 Emergent Literacy skills

Phonemic Awareness. Phonemic awareness includes the ability to identify sounds in words, to separate words into sounds, and to manipulate those sounds. The endline questionnaire tested phonemic awareness through identification of onset and rhyme sounds (first and last sounds, respectively).

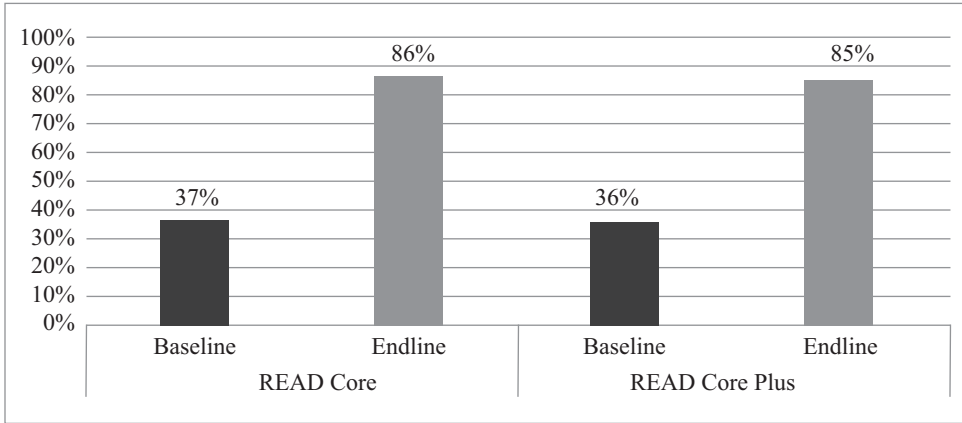
Similar Beginning Sounds. For this test, the assessor reads aloud three words, asking the student to identify the pair of words with similar beginning (onset) sounds. From regression, we see that similar to baseline, there is no significant difference between the Read Core and Read Core Plus group at the endline. But we can see that both the groups show a large improvement in performance from the baseline – more than 2.5 times. In the endline, each student could identify, on average, ninety percent pairs of similar beginning sound correctly. While at midline, Core Plus students performed better than Core students that is not the case in endline (Figure 2).

Figure 2: Percentage of Similar Beginning Sound Correct - Difference in Means between READ Core and READ Core Plus Students



Similar Ending Sounds. For this test, the assessor reads aloud three words, asking the student to identify the pair of words with similar ending sounds. Regression results are similar to similar ending sounds – no statistically significant difference between the two groups in the endline but a large increase in performance from the baseline by both the groups. In the endline, each student could identify, on average, eighty-five percent similar ending sounds correctly. Like similar beginning sounds, students from READ Core Plus performed poorly than the students from READ Core students in this subtask in the endline, though the result is not statistically significant (Figure 3).

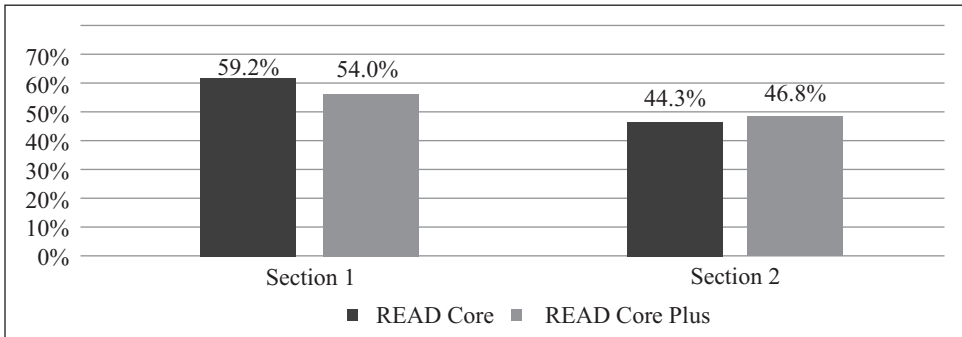
Figure 3: Percentage of Ending Rhyme in Words Correct - - Difference in Means between READ Core and READ Core Plus Students



Listening comprehension. In the listening comprehension subtask, non-reader student responds to questions about a comprehension read aloud to them. There are two stories for assessing listening or reading comprehension – section 1 or story 1 is relatively easier (equivalent to Grade II) with 80 words and section 2 or story 2 is relatively harder with 119 words (equivalent to Grade III). Literal questions (answers are available directly in the comprehension passage) and inference questions (not directly available in the comprehension passage) are included in the comprehension questionnaire. This test was conducted only for the endline.

Percentage of non-reader¹⁵ was quite low in the endline—below ten percent. But among those who are non-reader in endline, there is no statistically difference between Read Core and Read Core plus students in answering listening comprehension questions. Non-Readers in both groups perform better in story 1 compared to story 2; but this is understandable as story 1 is relatively easier compared to story 2 (Figure 4).

Figure 4: % Listening Comprehension Answered Correctly by Non-readers in Endline



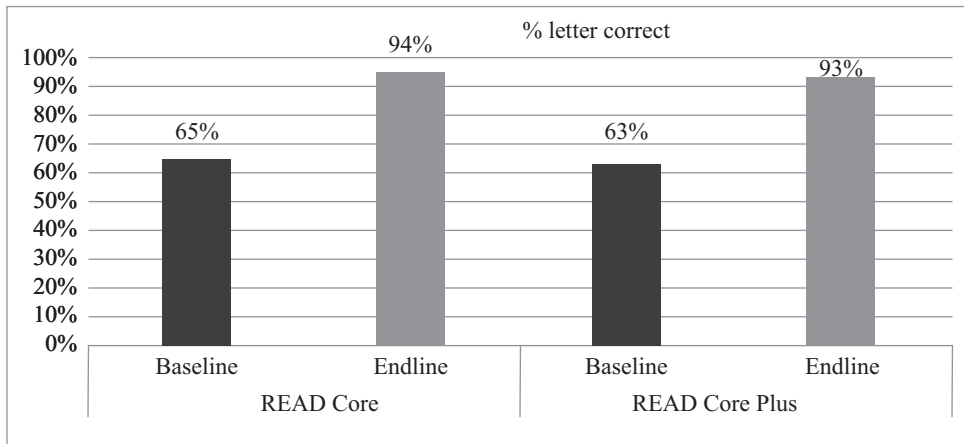
¹⁵ Details of readers and non-readers will be discussed in the following section.

6.3 Decoding

Letter Identification. In the letter identification subtask, students were asked to read aloud 50 letters of the Bangla alphabet. The full set of letters was listed in a random order to prevent reciting a memorized alphabet. This was done also to test actual automaticity of letter recognition and translation from print to sound.

We can see that there is no statistical difference between the Read Core and Read Core Plus group in endline performance. Both groups performed very well on this measure. This is expected as by Grade three, most students are expected to master the letter identification skill. On average, students from both the groups could identify 93% letters correctly (Figure 5)

Figure 5: Percentage letter correct -- Difference in Means between READ Core and READ Core Plus Students

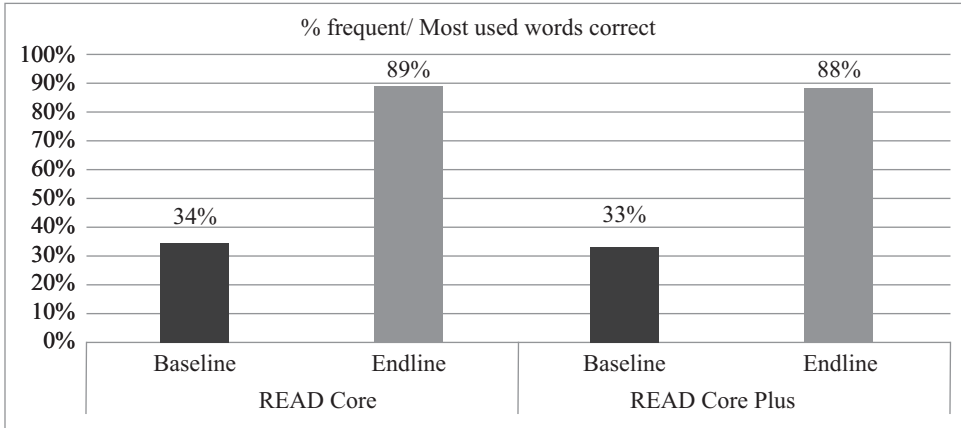


Familiar (most used) Word Reading. In the familiar word reading subtask, students were asked to read 20 words selected from the Bangla reading material. This subtask is meant for assessing sight recognition skills-decoding skills of the most frequent ('sight'¹⁶) words.

It can be seen that there is no significant difference in performance between the two groups in the endline. Yet again, we see a very large improvement in performance by both the group – more than 2.6 times from the baseline. On average, all students from both the groups answered more than 88% familiar words correctly. We also observe slightly better, though insignificant, performance by the students in the READ Core group (Figure 6).

¹⁶ Words that primary school students should recognize on sight, as many of these words are not easy to pronounce and thus must be memorized.

Figure 6: Percentage frequent/most used words correct-difference in means between READ Core and READ Core Plus students



Antonyms. Ten words, one at a time, were spoken out to students and were asked to say the antonym of each word. This test was performed to gauge students' vocabulary and their understanding of word meaning.

This test was performed only during the endline, and there is no significant difference between Read Core and Read Core Plus students and both groups performed well – they correctly answered 84.4% and 82.6% of the questions respectively.

Fluency and comprehension. Oral reading fluency is a measure of overall reading competence: the ability to translate letters into sounds, unify sounds into words, process connections, relate text to meaning, and make inferences to fill in missing information (Hasbrouck & Tindal, 2006). Using this complex process, skilled readers translate text into spoken language in a seemingly effortless manner (automaticity) and thus it can be used to characterize overall reading skill (Gove 2011).

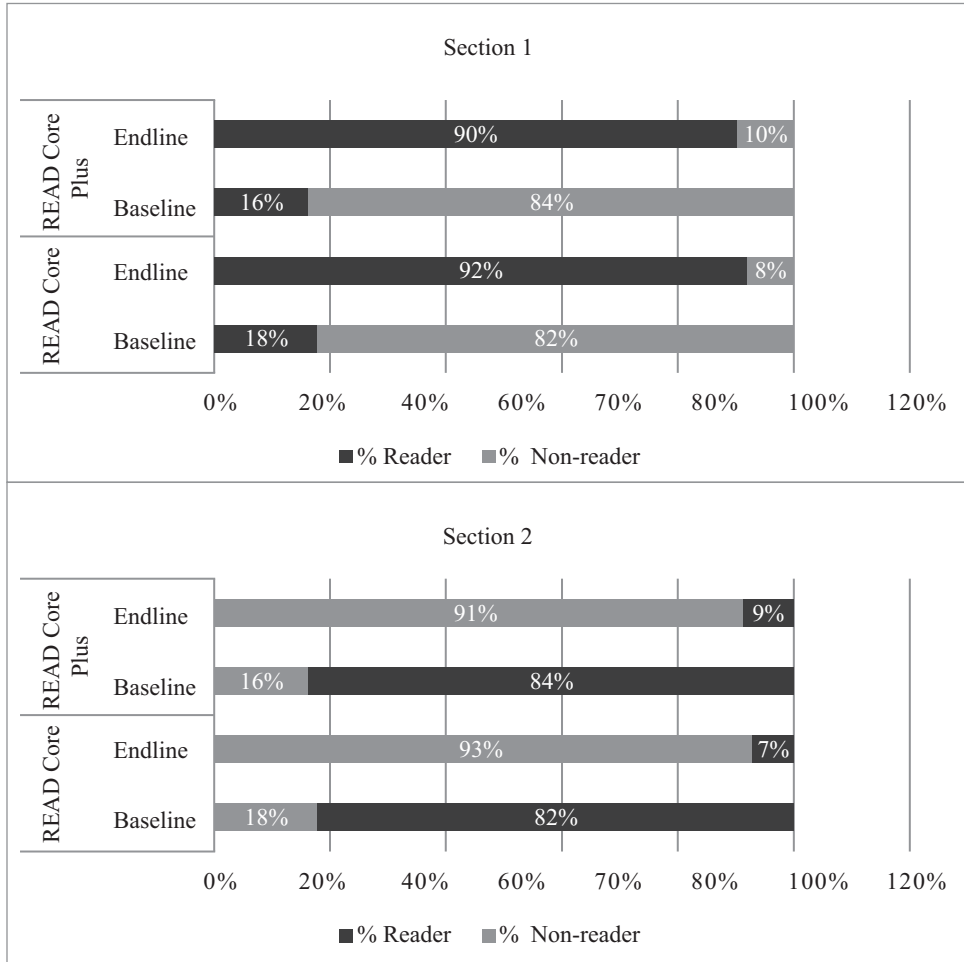
In the oral fluency subtask, students were asked to read two grade-appropriate comprehensions, developed by the SCI team. As mentioned earlier, the first comprehension subtask has a relatively easy story of 80 words (appropriate for Grade II) and the second comprehension subtask has a harder story with 119 words (appropriate for Grade III). For each story, this sub-task was administered only for those students who could read at least five words correctly in 30 seconds (reader) from that story.

Reader. Students who could read at least five words in 30 seconds were marked as readers. This criterion for being a reader was developed by SCI - READ team¹⁷. There are two reading comprehension sections, one with a lower difficulty level is from Grade II reading text (Sec 1) and another one's difficulty level is of Grade III level (Sec 2). Again, we find no significant

¹⁷ EGRA toolkit suggests that a student should be able to read enough text for the first question to be considered as reader. Here, 30 seconds was enough time for the students to read text to answer first questions as suggested by RTI.

difference between the Read Core and Read Core Plus students on this measure. For both stories, more than 90% students in both the groups are self-reliant readers in the endline. Gain in performance is almost 5 times higher than the baseline (Figure 7).

Figure 7: % of Students Who are Self-reliant Readers (Section 1 and 2) - Difference in Means between READ Core and READ Core Plus Students

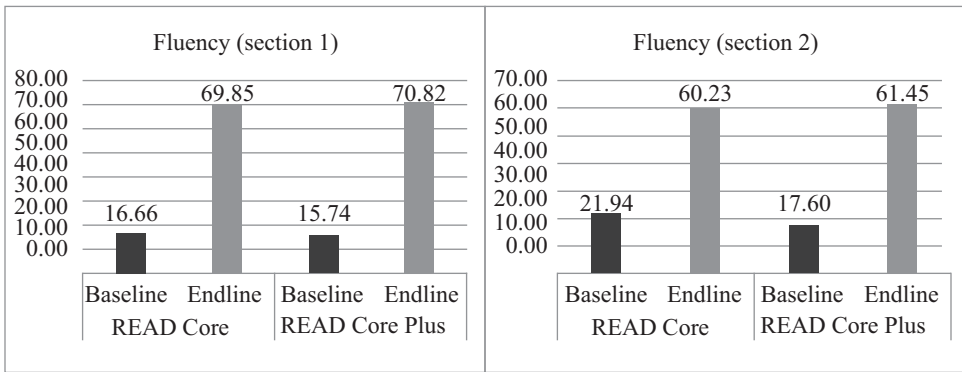


6.4 Oral Reading Fluency (ORF)

ORF was not administered separately and was calculated from the reading comprehension subtask. ORF was measured by calculating percent of words, on average, a student could read correctly in 60 seconds.

There was no significant difference, (based on regression not shown here), between Read Core and Read Core Plus students in reading words correctly in 60 second, but again the improvement of both the groups is high from baseline. In both the stories, students from both groups could read more than 60 words per minute. Though, in reading five words in 30 seconds (reader), READ Core students have done better than READ Core Plus students, the trend seems to be slightly different in fluency, but the difference was not found to be statistically significant (Figure 8).

Figure 8: Correct Number of Words Per Minute - Difference in Means between READ Core and READ Core Plus Students



Accuracy. Accuracy was also not administered separately and was calculated from the reading comprehension subtask. Accuracy was measured by calculating percent of words read correctly. We can see from the below table that there is no statistically significant difference between Read Core and Read Core Plus students in the endline for both the stories, both the groups read more than 95% words correctly in the endline. This is a large gain from the baseline.

6.5 Comprehension

Students were asked 10 comprehension questions that included seven literal questions and three inference questions. Readers in both READ Core and READ Core Plus group have done very well in answering comprehension questions from both the sections in the endline, and there is no significant difference in performance between the two groups. Overall, on average, readers from both groups could answer more than 80% comprehension questions correctly in both sections (See Table 4).

Reader with comprehension questions. This test was not administered separately and was calculated from the comprehension questions subtask. A binary variable (0 or 1) was generated indicating students who answered eight or more comprehension questions correctly and who answered more than eight comprehension questions correctly.

We do not find any significant difference between the Read Core and Read Core Plus students on these measure in the endline. Similar to most other sub-tasks, both groups have improved by a very large margin in the endline compared to the baseline. In the endline, more than 85% students answered at least 5 comprehension questions correctly in both sections.

Table 4: Comprehension -difference in means between READ Core and READ Core Plus students

Variable Name	Endline				Baseline			
	READ Core	READ Core Plus	p-value	Sample Size	READ Core	READ Core Plus	p-value	Sample Size
Accuracy score for Readers (Section 1)	96.4%	97.1%	0.229	1577	87.2%	87.3%	0.945	397
% of comprehension questions, correctly by readers (Section 1)	88.6%	89.0%	0.744	1577	57.0%	58.0%	0.724	397
% who answered 8 or more comprehensions correctly (all students) (Section 1)	77.7%	74.6%	0.377	1732	4.8%	4.3%	0.680	2319
% of students who answered 5 or more comprehensions correctly(all students) (Section 1)	89.1%	87.6%	0.602	1731	12%	10.7%	0.776	1351
Accuracy score for Readers (Section 2)	95.1%	95.5%	0.539	1407	73.6%	73.4%	0.955	397
% of comprehension questions, correctly by readers (Section 2)	81.6%	80.6%	0.671	1529	7%	6%	0.391	2319
% who answered 8 or more comprehensions correctly (all students) (Section 2)	71.4%	66.4%	0.257	1529	2.1%	1.1%	0.158	2319
% who answered 5 or more comprehensions correctly (all students) (Section 2)	86.4%	84.4%	0.490	1528	7.7%	6.6%	0.392	1676

Overall, we do not see any significant difference in performance between the Read Core and Read Core Plus students in the endline. Importantly, in both the groups 90% students are reader and the readers, on average, answered more than 80% comprehension questions correctly with fluency of more than 60 words per minute.

7. Heterogeneity Analysis

7.1 Difference in Endline Performance by Student Background Characteristics

Using the coefficients from the multi-level regression models, we tried to identify the relation between reading skills and students' background information such as gender, socio-economic status (SEC), attendance in Early Childhood Development (ECD) programs and

availability of reading materials at home. Table 5 presents a simplified summary of our findings. We see that boys (compared to girls) and students in lower SES (compared to those in higher SEC) are significantly falling behind in fluency in both the comprehension sections, similar beginning sounds and similar ending sounds ($p < 0.05$). The trend was similar in baseline and midline as well. Also, students with fewer reading material at home are significantly falling behind in most of the reading skill tests than the students with more reading materials at home. The trend was similar in baseline, interestingly there was no relation between available reading materials at home and reading tests at midline. Students who did not attend any ECD are falling behind compared to those who attended ECD in letter identification and most used words at baseline, but the advantages was not seen in the midline. It is possible students who have fewer reading materials at home and did not attend ECD managed to catch up with those with more books and ECD attendance by Grade II. But why they fell back again is a matter of further investigation.

Table 6: Who is Falling Behind in the Endline-by background?

Endline			
	Gender	Socio-Economic Status	Reading material at home
Similar beginning sounds	Boys		Few reading materials
Similar ending sounds	Boys		Few reading materials
Letter identification	Boys		Few reading materials
Most used words	Boys	Low SES	Few reading materials
Antonyms	-	Low SES	Few reading materials
Section 1 Grade 2 level text			
Reader		Low SES	Few reading materials
Accuracy	Boys	Low SES	Few reading materials
Fluency	Boys	Low SES	Few reading materials
Reading Comprehension		-	Few reading materials
Reading Comprehension with more than 8 questions	-	Low SES	Few reading materials
Section 2 Grade 3 level text			
Reader	Boys	Low SES	Few reading materials
Accuracy			Few reading materials
Fluency	Boys	Low SES	Few reading materials
Reading Comprehension		Low SES	Few reading materials
Reading Comprehension with more than 8 questions	-	Low SES	Few reading materials

Note: (-) indicates no difference

7.2 Difference in Endline Performance by Division¹⁸

We investigate whether there is a significant difference in endline performance by division, and we find that compared to Dhaka, Rangpur has performed significantly better in all skill

¹⁸ We should be careful while interpreting these results due to very small sample size. Implication of very small sample size is low power for generalizing the results.

areas – phonemic awareness, decoding and comprehension. Students in Rangpur also performed better than all other Divisions in most of these subtasks.

In phonemic awareness, Rangpur exhibited highest level of performance, answering correctly between 95% and 97% similar beginning sounds, which is significantly better than Dhaka ($p < 0.01$). They exhibit best performance, answering correctly 92% to 94% similar ending sounds, significantly better than Dhaka ($p < 0.01$)

In decoding skills, students from Rangpur exhibited highest level of performance answering correctly 92% to 97% letters, significantly better than Dhaka ($p < 0.01$). They exhibit highest level of performance, correcting 94% to 97% most frequent words, significantly better than Dhaka (significant at $p < 0.05$). We observe similar performance of students in Rangpur in answering antonyms, correcting 87% to 93% words correctly, significantly better than Dhaka ($p < 0.05$).

Students in Rangpur have also performed significantly better in Comprehension skills as well. More than 98% students in Rangpur are reader, significantly better than Dhaka (significant at $p < 0.1$). Students from Rangpur also performed significantly better compared to Dhaka in correctly answering comprehension questions in both section 1 ($p < 0.01$) and section 2 ($p < 0.01$). They are also the best performer in the two subtasks, answering between 94 and 96% comprehension questions correctly in section 1 and 88 to 92% in section 2.

We also observe, students in Dhaka performs worse than other divisions in almost all the subtasks where there is a significant difference. We also find that the students in Sylhet division have performed worse than Dhaka in most of the subtasks, though significantly worse only in two tasks. Students in Chittagong and Barisal have also done worse than Dhaka on many subtasks.

7.3 Difference in Endline Performance by Student Baseline Performance

Here we examine whether there is a difference in endline performance by how the students performed in the baseline, in other words, whether there is a differential performance between the weaker or stronger students at the baseline. In order to do this analysis, we categorized the students in four groups based on their baseline performance from best to worst.

When we look at just the endline performance, we see that students in both Core and Core Plus group with higher baseline performance performs significantly better than those with lowest baseline performance. What is interesting is that students with lowest baseline performance in Core Plus group gained proportionately more compared to those with higher baseline performance. This relatively strong gain also applies to both groups, with no advantage for core plus participants.

Within the READ Core Plus group, we also categorized the students by their attendance in reading camp in two groups – one with regular attendance and the other with irregular attendance. When we analyze the endline performance of Core Plus students by attendance and by performance in baseline, we see that there is no difference between the lowest performing students who regularly attended the reading camp and those who did not. If we look at the gain in endline performance from baseline, again we see that there is almost no

statistically significant difference between the regular and irregular attendees with lowest baseline performance. These findings suggest that the overall gain of the Core Plus students with lowest baseline performance cannot be attributed to their participation in the reading camp. Additionally, we see that irregular and regular CRC attendees with better baseline score gained significantly less compared to the irregular attendees with the lowest baseline performance.

8. Discussion and Conclusions

This is a randomized controlled clustered study to estimate the impact of community-based intervention (READ Core Plus) over a school-based intervention (READ Core). From the baseline analysis, we find that groups were balanced before the program intervention. Also, from our attrition analysis, we found there is no significant difference in most of the baseline characteristics between re-contacted READ Core and READ Core plus students and between students who were lost from both READ core and Core Plus interventions for various reasons.

A critical question is why the additional intervention through the reading camp did not make a significance difference in student's achievement in reading skills.

We investigated whether regular attendance in reading camp had an additional impact on reading skills. When we compared the endline scores of READ Core students with the scores of Core Plus regular community reading camp attendees, we did not see any significant difference. Further, we find that students with low baseline score gained proportionately more than better performing students in both Core and Core Plus, but among the lowest performers in Core Plus, there is no statistical difference between regular and irregular reading camp attendees; this reconfirms the finding that no significant improvement of performance resulted from the reading camp.

Could the frequency of exposure – how many days a week a reading camp is held -be a factor? In the well-known study by Banerjee, et al.,¹⁹ the researchers found that an intervention in India with trained volunteers to teach children to read outside public schools had a large impact improving the reading skills of the students who attended the camps. This model is quite similar to the reading camp model. But the intensity and durations are different. In their study, the typical “reading course” lasted two to three months, with classes held every day outside of school. In the CRC model, each student attends the camp once a week though it lasts longer – about eight months. The frequency and intensity of camps, rather than the duration may be more important for improving reading skills of students.

Gain in reading skill tests, as presented in this report, shows large improvements in all the components of reading skills. These large improvements, particularly in comprehension questions in this RCT study, shows a contrast with previous findings in similar evaluations by Save the Children. For example, in the endline evaluation of the Newly Nationalized

¹⁹ Banerjee, Abhijit et al., “Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India,” Working Paper (National Bureau of Economic Research, September 2008), <http://www.nber.org/papers/w14311>.

Primary Schools (NNPS),²⁰ we see READ Core readers from both Grade II and Grade III students could answer only about 38% comprehension questions correctly, whereas in this study all readers answered more than 80% comprehension questions correctly.

There may be several explanations for this large improvement. First, in this study, we followed the same students from Grade I to Grade III, ensuring that they received the program for at least 2.5 years. On the other hand, in the endline with NNPS, we followed same schools but not the same students. Thus, it is quite likely that there was a significant number of new students who received the program for a much shorter duration. The duration of the program as well as the intensity and quality of the activities may be important in enhancing comprehension ability and reading skills in general.

Provision of school facilities, general teaching-learning environment and readiness of students for school may matter in producing an impact. NNPS schools were much worse than RCT schools in terms of infrastructure, teacher quality, students' background and these differences may have an influence on the results. We may check the baseline characteristics of NNPS and RCT schools as well as students to verify this hypothesis.

We also observe that there are significant regional variations in student achievement in different components of reading skills. Students from Rangpur performed significantly better than the students in other divisions in most of the tests: letter identification, most used words, similar beginning sounds, similar ending sounds, antonyms, fluency and accuracy. It is important to note that Khulna has performed lowest in most of the emergent and literacy skills, but that is not the case for reading comprehension. Dhaka also performed worse than other divisions in most of the subtasks. A qualitative assessment of the program implementation in Khulna and Dhaka region might give us a clearer idea in identifying the reasons for such results.

Finally, boys are falling behind in most of the emergent and decoding skills compared to girls. It is also important to note that students with fewer reading materials at home are falling behind in almost all the reading tests.

In recapitulating the findings and drawing conclusions, the following may be highlighted.

- The percentage of self-reliant readers have increased by about five times during the intervention period. This is partly the result of maturation of students during this period, and there has been no control group to isolate the effect of the interventions. However, it should be noted that a similar overall progress in reading skills has not happened with primary school students in the country as seen from National Student Assessment results.
- The school-based core interventions have produced or at least contributed to the positive results; the additional community reading camp intervention has not added significantly to student achievement. This suggests the importance of school-based activities and cost-effectiveness of these classroom activities, because additional interventions in the community inevitably mean additional cost and management efforts.

²⁰ This was not an RCT.

- Although, students coming from relatively privileged home background (one indication of it is more reading materials at home), gained relatively more from the interventions, those with disadvantaged home background also had gained considerably in different reading sub-tasks when compared with their baseline skills. This finding suggests need for examining community-based interventions which may be specifically geared to the students with poor home environment.
- Participation in ECD/preschool made no difference in later reading skills achievement of students participating in the program. This is a commentary on the design and quality of implementation of the preschool program which is now being offered to most primary school entrants.
- Girls in general have shown better performance in reading skills. While everything should be done to overcome pervasive gender inequality in society, reverse disparity affecting boys is not acceptable. The causes and circumstances of this phenomenon need to be examined and remedied.

Finally, the vital policy and strategy issue is the recognition in primary education development programme of Bangladesh the critical importance of enabling students to acquire the foundational skills of reading with comprehension and, therefore, organizing and designing teaching-learning for this purpose. Otherwise, the approach and methodology of early grade reading that have shown good results cannot be mainstreamed into the primary schools.

References

- Annie E. Casey Foundation (2010). "Early Warning: Why Reading By the End of Third Grade Matters." Baltimore, MD, USA: Annie E. Casey Foundation.
- Banerjee, A. et al. (2008). "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India," Working Paper (National Bureau of Economic Research, September 2008), <http://www.nber.org/papers/w14311>.
- Crouch, L.(2012). "Why Early Grade Reading: An Economist's Perspective. Presentation given at 'All Children Reading Workshop'." Kigali, Rwanda, 28 February 2012.
- Cunningham, A. E.; Stanovich, K. E. (1997). "Early reading acquisition and its relation to reading experience and ability 10 years later." *Developmental Psychology*, Vol 33(6), Nov 1997, 934-945. <http://dx.doi.org/10.1037/0012-1649.33.6.934>
- Directorate of Primary Education, Government of Bangladesh (2012). National Student Assessment for Grades 3 and 5 National Report, 2011. Government of the People's Republic of Bangladesh Directorate of Primary Education, 29 December 2012.
- Directorate of Primary Education, Government of Bangladesh (2014). National Student Assessment for Grades 3 and 5 National Report, 2013. Government of the People's Republic of Bangladesh Directorate of Primary Education., 18 December 2014.
- Hernandez, D. (2011). "Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation." New York, USA: The Annie E. Casey Foundation
- Spira, E. G., Bracken, S. S., Fischel, J. E.(2005). "Predicting Improvement after First-Grade Reading Difficulties: The Effects of Oral Language, Emergent Literacy, and Behavior Skills." *Developmental Psychology*, Vol 41(1), Jan 2005, 225-234. <http://dx.doi.org/10.1037/0012-1649.41.1.225>