I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15920

Delivering Remote Learning Using a Low-Tech Solution: Evidence from a Randomized Controlled Trial in Bangladesh

Liang Choon Wang
Michael Vlassopoulos
Asad Islam
Hashibul Hassan

FEBRUARY 2023

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15920

Delivering Remote Learning Using a Low-Tech Solution: Evidence from a Randomized Controlled Trial in Bangladesh

Liang Choon Wang
Monash University

Asad Islam
CDES and Monash University

Michael Vlassopoulos
University of Southampton and IZA

Hashibul Hassan
Jagannath University

FEBRUARY 2023

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Delivering Remote Learning Using a Low-Tech Solution: Evidence from a Randomized Controlled Trial in Bangladesh*

The Covid-19 pandemic caused prolonged school closures worldwide. Children in resource-poor settings were particularly affected given their limited access to remedial distance learning opportunities through the internet, television, and radio. To address the poor access to formal education, we designed an educational intervention consisting of a set of audio lessons that were delivered through mobile phones to primary school students using Interactive Voice Response (IVR). During the 15-week program period, parents could access the lessons for free by calling a designated phone number and listening to a lesson with their child at any time. We delivered the randomized intervention to 1,763 primary school children across 90 villages in Bangladesh during the 2021 Covid-19 school closures. The intervention improved the test scores of children in literacy and numeracy by 0.60 Standard Deviations (SD). Additionally, the intervention led to an increase in the amount of time that parents spent on homeschooling. The intervention was particularly beneficial for academically weaker students, those from the poorest strata, and those with less-educated caregivers. Our results suggest that this scalable and low-cost intervention could be leveraged in similar settings to address learning losses of marginalized students.

JEL Classification: C93, I21, I24

Keywords: school closures, remote education, Interactive Voice Response (IVR), COVID-19, randomized controlled trial, Bangladesh

Corresponding author:
Asad Islam
Department of Economics
Monash University
Victoria 3800
Australia

E-mail: asadul.islam@Monash.edu

* Authors' names appear in reverse alphabetical order. We thank Buly Cardak, Gaurav Datt, Paul Glewwe, Claudio Labanca, Albert Park, Abu Siddique, Agne Suziedelyte, Rebecca Thornton, seminar and conference participants at Asian Development Bank (ADB), Curtin University, Monash University, Bangladesh Institute of Development Studies, BRAC Institute of Governance and Development (BIGD). This project is registered at the AEA RCT registry: AEARCTR-0007931; ethical clearance for this project was received from Monash University, Australia (Project no. 27474). This project is funded by an EdTech Hub Covid-19 research grant.

1 Introduction

The Covid-19 pandemic affected the lives of billions of people around the world in numerous ways, with disruptions in education being a key domain. More than 1.5 billion students of all ages across 180 countries have been impacted by the closures of educational institutions (UNESCO, 2022). Mounting evidence indicates that school closures have led to large learning losses worldwide, especially for children from disadvantaged backgrounds and those living in low and middle-income countries (Agostinelli, Doepke, Sorrenti, & Zilibotti, 2022; Engzell, Frey, & Verhagen, 2021; Moscoviz & Evans, 2022; Patrinos, 2022; Singh, Romero, & Muralidharan, 2022). These children are more vulnerable to disruptions in formal, in-person education, as they have limited access to distance learning resources and may lack adequate parental support for their learning (Rahman & Sharma, 2021; UNICEF, 2020). Therefore, there are growing concerns that school closures will exacerbate pre-existing education inequalities. These concerns highlight the need for low cost and effective remote learning solutions that can be mobilized when schools are forced to close due to public health emergencies or when other causes, such as natural disasters, wars, strikes, and political unrests, trigger educational disruptions.

Bangladesh provides a good setting to study remote learning interventions. The country experienced one of the world's longest periods of school closures during the Covid-19 pandemic, with around 37 million children having had their learning disrupted (UNESCO, 2022). A rapid survey by the World Bank found that only around 40% of students had access to remote learning in the first few months of the Covid-19 pandemic (Biswas et al., 2020). Even after a year of school closures, more than 40% students did not have access to remote learning (Rahman et al., 2021). Only a small percentage (5%) of children aged 5-15 years had access to a computer, and the active internet usage rate (28.8%) was also low due to the lack of compatible devices and high data costs (DataReportal, 2021; Rahman et al., 2021). More than half of the school-aged children did not have access to TV, and most of the students who had access did not benefit from the TV and online lessons that were available during school closures. The situation appears particularly grim for students in rural and disadvantaged areas (Beam, Mukherjee, Navarro-Sola, Ferdosh, & Sarwar, 2021; Hassan, Islam, Siddique, & Wang, 2021).

This paper reports evidence from a feature phone-based remote learning intervention aimed at addressing the learning needs of children during the Covid-19 pandemic in a resource-constrained context. Our educational program was delivered during school closures in

Bangladesh, targeting over 1700 primary school children across 90 villages in two southwestern districts (Khulna and Satkhira) of the country. The education program in question involved the delivery of pre-recorded audio lessons by using Interactive Voice Response (IVR) technology.[1] A key advantage of this technology is that it provides flexibility regarding learning levels and the timing of learning: lessons of different proficiency levels can be stored in a telecom server, and learners can choose lessons at their competence level, and when to access them without having to follow a pre-arranged schedule. More than 96% of households in rural Bangladesh have access to a mobile phone, but less than one-third of them have smart phones (Hassan et al., 2021), prompting us to offer a resource that could be accessed using a simple feature phone. In designing the audio lessons, we employed the distance learning method of Interactive Audio Instruction (IAI), in which learning content is delivered through pre-recorded audio broadcasts and learners engage actively through questions and exercises (Bosch, 1997). This method was originally conceived to deliver lessons through radio, and has been shown to be effective in improving learning outcomes in conventional classrooms (Anzalone & Bosch, 2005; Ho & Thukral, 2009).

The educational program covered two main areas – literacy and numeracy – and was divided into 60 audio lessons. The lessons were delivered over a 15-week period, with each lesson lasting between 16 and 18 minutes. The audio lessons were structured as pre-recorded conversations among four characters, two teachers and two students, following the IAI methodology. During the lesson, students were asked to engage in some activities following the teachers' instructions, such as clapping, standing up, counting, etc. The teachers used regular pauses and cues as well as played music and songs during the recorded lessons so that the listeners (program children) could complete similar tasks with the help of their caregivers, typically their mothers. The caregivers could select and access any lesson at any time for their child during the program period, as we did not impose that any specific sequence be followed. To deliver the audio lessons, we established two IVR-enabled toll-free numbers.

Beyond numeracy and literacy, we were also interested in investigating whether children's leadership skills could be improved through this phone-based remote learning method. Generally, leadership is considered as a complex, multidimensional advanced competency rather than a fixed, genetic personality trait (Karagianni & Jude Montgomery, 2018). Thus, it

[1] Interactive Voice Response (IVR) is an automated phone system that allows humans to access information via a touch tone system or a voice response system.

is perceived as a dynamic skill that can be developed by appropriate interventions (Sisk, 1993). Evidence also indicates that young people who received various leadership interventions improve on a range of outcomes (Karagianni & Jude Montgomery, 2018). Therefore, in a separate treatment group of students, we offered an additional leadership skill module adapted from the Lead Africa program (LEAD, 2021). The contents were designed to engage students in various non-academic activities with their adult supervisor at home. The lesson plan covered a broad range of leadership skills, such as communication, planning, patience, empathy, sympathy, compassion, and perseverance, and encouraged extended interaction between caregivers and children at home. Furthermore, as schools were closed and social interactions with other children were limited, in addition to numeracy, literacy and leadership skills, we also examined whether the intervention had any effects on children's noncognitive skills and behavioral difficulties. For example, lockdown and limited interactions with other children could exacerbate children's behavioral problems, such as tantrums, nervousness, lack of control over emotions, and hyperactivity. The various modules that were offered could help children improve their noncognitive skills and decrease their behavioral difficulties.

We implemented this education program in a three-arm clustered Randomized Controlled Trial (RCT) across 90 villages (30 villages in each arm). In the *Standard* group, we offered participating children the literacy and numeracy module; in the *Extended* group, we offered the leadership module in addition to the literacy and numeracy module; and in the Control group, no intervention was offered. One challenge faced by remote learning educational programs is that participants might not engage with the material either because they face financial or time constraints or because they do not find the content appealing. To assess this possibility, we monitored the take-up rate and usage of the learning material throughout the intervention. The data show that participants engaged substantially with the lessons, as more than 70% of participating children completed at least two-thirds of the lessons.

We find that the IVR-based educational program led to substantial improvements in the learning outcomes of children, as measured by assessment tests on literacy and numeracy that we administered in the endline. Relative to children in the Control group, treated children in the Standard and Extended groups experienced a 0.60 standard deviation (SD) and 0.63 SD improvement (about 30% higher total test score), respectively, on the total test score that combines the two tests. Importantly, we find that the intervention was especially beneficial for academically weaker students, students from the poorest strata, and students with less-educated

caregivers. Thus, the intervention is particularly valuable for more vulnerable groups of students and could contribute toward reducing educational inequalities.

We find weak evidence that children in the Standard treatment experienced improvements in behavioral issues: they showed less emotional symptoms, conduct problems, hyperactivity, peer problems and a more prosocial attitude. On the other hand, we do not find evidence that the intervention improved the leadership and communication skills of children, which was the aim of the extra module included in the Extended group. We also do not find significant impacts on a range of noncognitive skills such as impulsivity, grit, growth mindset, and empathy. These findings suggest that fostering noncognitive skills may require larger investments than this type of short remote learning programs can provide.

Overall, our findings suggest that the phone-based educational program can be an effective and scalable distance learning tool to improve learning outcomes of students in a context with no access to formal in-person education and limited access to alternative content through digital technological devices. An important aspect of the intervention in this study is its wide accessibility, as audio lessons can be accessed via basic feature phones, which are widely available even in underprivileged families. A second feature is scalability. The cost of this 15-week intervention was USD 27.5 per student, of which USD 13.2 were variable costs and USD 14.3 were fixed costs. Scaling up the program would likely further reduce the per-student cost considerably.

Our paper contributes to a recent literature that investigates alternative programs that address educational disruptions, such as hiring paid instructors or volunteer tutors to help students with their learning over the phone (Angrist, Bergman, & Matsheng, 2022; Crawfurd, Evans, Hares, & Sandefur, 2021; Hassan et al., 2021; Lichand & Christen, 2021; Schueler & Rodriguez-Segura, 2022).[2] One challenge facing these programs is that it can be challenging to scale them up, especially in a low-income country context, as paid instructors are costly and volunteers are difficult to retain (Islam, Malek, Tasneem, & Wang, 2022). Also, the feature of flexible study hours is absent from other feature phone-based educational programs. The flexible delivery method of our program accommodates the resource constraints rural households

[2] Research on feature-phone-based educational interventions is not a new area of interest. In the last decade, excluding the Covid-19 pandemic period, various studies have demonstrated the effectiveness of this platform. However, these studies have mostly been restricted to teacher–caregiver rather than teacher-student engagement (Bergman, 2021; Berlinski, Busso, Dinkelman, & Martinez, 2016; Hurwitz, Lauricella, Hanson, Raden, & Wartella, 2015; Matthew A Kraft & Dougherty, 2013; Mayer, Kalil, Oreopoulos, & Gallegos, 2015), or high school graduates (Bird et al., 2021; Castleman & Meyer, 2020), or adult learners (Aker & Ksoll, 2019; Ksoll, Aker, Miller, Perez, & Smalley, 2015), rather than primary graders.

typically face, such as having only one phone in each household, and unstable mobile networks and electricity supply.

The main contribution of the current paper is to provide evidence of the effectiveness of a remote learning intervention delivered during school closures via IVR. This approach could be a viable alternative to more human resource intensive programs, as it is scalable and offers some flexibility regarding learning levels and learning delivery schedules to learners and their caregivers.

2 Intervention and Research Design

2.1 Background

When all schools were closed due to the Covid-19 pandemic in March 2020, the government of Bangladesh responded to the ensuing educational crisis by providing multimodal distance learning. By the first week of April 2020, the Ministry of Primary and Mass Education (MoPME) and the Ministry of Education (MoE) started remote learning through asynchronous classes broadcast via national television and online platforms (Rahman & Sharma, 2021). Online resources were already developed and available on various sites but expanded during the school closures. Radio broadcasting was added later in the year.

Despite the quick delivery of multimodal distance learning, a rapid survey by the World Bank found that only around 40% of students had access to remote learning in the first few months of the Covid-19 pandemic (Biswas et al., 2020). Even after a year, a significant portion of children remained outside of distance learning coverage; 44% and 36% of rural households and urban-slum households did not have access, respectively (Rahman et al., 2021). This lack of access led to poor learning outcomes as only 18% of primary graders and 38% of secondary graders were actively learning through assignments in August 2021 (Rahman et al., 2021).[3]

2.2 The Intervention

The aim of the intervention was to deliver interactive audio content via IVR to improve the learning of primary-school students. We next explain the main features of the intervention.

[3] MoPME and MoE gave various assignments via distance education programs and online platforms to engage students in learning and to assess them.

2.2.1 IVR-based Education

Interactive Voice Response (IVR) is an automated phone system technology that allows incoming or receiving callers to access information by traversing or navigating a pre-designed flow. Navigation to different points of the flow can be done by either voice command or keypad selection by the caller. Once the caller makes a call to or receives a call from an IVR-enabled number, there is no need for any human agent. A pre-recorded message can guide them to the desired landing node with preferred information. Though both radio and IVR platforms only support audio lessons, IVR has an important advantage over the radio: listeners can select lessons, i.e., there is no fixed broadcasting sequence.

2.2.2 Interactive Audio Instruction (IAI)

The original design of Interactive Audio Instruction was created back in 1970 as Interactive Radio Instructions (IRI) to teach mathematics via radio in Nicaragua (Bosch, 1997). IAI is an instructional approach that turns a one-way technology into a tool for active learning as it requires learners to stop and react to questions and exercises through verbal response, to engage in group work, and physical and intellectual activities while the program is on the air (Bosch, Rhodes, & Kariuki, 2002). Facilitators play an important role in supervising the progression of the lessons (Ho & Thukral, 2009).

Just before the intervention, the field staff of the Global Development and Research Initiative (GDRI), our local partner, visited the sample households to provide a guidebook and briefing on the IAI method. In particular, they explained how the interaction would happen, what would be the role of caregivers, and what the caregivers would need during the lesson play. Caregivers were the facilitators in our intervention. For example, they were invited to draw a few figures, show some letters, or write numbers to engage with their learners during the IAI sessions played over the IVR. In Appendix A we provide a few examples of such activities.

2.2.3 Content of Intervention

The 15-week intervention included three elementary educational modules – literacy, numeracy, and leadership – divided into 75 audio lessons, with each lesson lasting between 16 and 18 minutes (Table A1-Table A3 in the Appendix A provide more details of these modules). Caregivers accessed these pre-recorded audio lessons via Interactive Voice Response (IVR) by first calling a toll-free number. Caregivers could choose and access any lesson at any time for the child participants, without following any specific sequence in the curriculum. Figure B1and

Figure B2 in the Appendix B provide an illustration of the IVR journey experienced by the participating caregiver–child pairs.

The audio lessons were pre-recorded conversations among four characters: two teachers and two students. During the conversation, students were asked to do some activities according to the teachers' instructions, such as clapping, standing up, counting, making plans, etc. The teachers used regular pauses and cues, as well as playing music and songs during the recorded lessons so that the program children could complete similar tasks with the help of their caregivers (mothers in 78% of cases) or any other adult members of the household (in 22% of cases).

We designed the modules for this program with the support of two international organizations and a group of local curriculum experts. Overall, these modules were developed to supplement the national curriculum and support learning in household settings. Section A-3 of the Appendix A provides detail content and lesson plan.

2.3 Treatments

We randomized each of the 90 participating villages into one of three groups (see Figure 1 for a summary of the research design):

- In the *Standard* treatment group, we offered the literacy and numeracy modules (60 lessons) to households in 30 villages.
- In the *Extended* treatment group, we offered the leadership module in addition to the literacy and numeracy modules to households in 30 villages. The leadership module focuses on development of leadership, listening and communication, and planning (see Section A-3.3).
- In the *Control* group (30 villages), no intervention was offered.

Five lessons (two literacy, two numeracy and one leadership) were offered each week. After each lesson, a quiz was played to the listener, and the answers were recorded in the IVR flow. To motivate listeners to complete the lesson, 15 listeners were randomly selected each week from the pool of listeners who answered quizzes correctly during that week to receive USD 3 as a prize via mobile financial services (MFS).

Caregivers could choose and access any lesson at any time for their child during the program period. We did not mandate any fixed sequence in our curriculum, i.e., if any learner found any lesson easy, they could skip the lesson and move to the next one. This is because students

usually learn better if they are provided with educational content that matches their level (Banerjee, Cole, Duflo, & Linden, 2007). Thus, by offering a menu of choices participating children could adjust their learning according to their competencies.

3 Data and Empirical Method

3.1 Sample

In partnership with the Global Development and Research Initiative (GDRI) — a local non-governmental organization (NGO) in Bangladesh, we worked with 1,763 primary school-aged children and their caregivers in 90 villages in two southwestern districts (Khulna and Satkhira) (see Figure 2). GDRI had worked with a larger sample of children (more than 7,500) across 223 villages in these areas before the Covid-19 pandemic.[4] We, therefore, had access to household contact information and pre-pandemic learning levels of the children.

From the list of contacts provided by GDRI, we, first, randomly selected 90 villages and then 3,000 households with mobile phone numbers. We were able to reach and complete a baseline survey for 2,400 children from 2,387 households in May 2021 (see Figure 3 for the project timeline). Others did not respond, or the phone was inactive, invalid, or switched off, or they were not interested in this intervention. We randomly selected about 16–22 children[5] from each of these 90 villages. Our final sample size was 1,763 children from 1,755 households at the baseline. At the endline, we reached 1,687 households for the endline survey and assessments. Survey attrition rates are not different across the treatment arms (Table B1 and Table B2 in the Appendix B).

3.2 Data Collection

Baseline assessment tests were conducted a year before this intervention as a part of GDRI's activities in the area. Due to the Covid-19 related school closures and mobility restrictions during the baseline period, we did not reassess the children immediately before the intervention commenced. During the intervention, we also collected IVR-flow usage data, i.e., duration of lesson-play by each participant, access time and date, frequency, etc. from the server. After the 15-week intervention, we surveyed the households and children again.

[4] These children were part of a completed project named Investing in our future by GDRI.
[5] There are four villages with less than 16 children: 10 children each in 2 villages, 11 children in 1 village and 13 children in 1 village. We capped the sample size at 22 children per village to match the budgetary allocation of this study.

Both the baseline and endline surveys contain information related to demographics, income, employment status, household asset composition, livelihood, caregiver involvement in education, the children's educational situation, and the households' private educational investment. At the endline, a team of two members — one assessor and one enumerator — visited each household. The assessor conducted the literacy and numeracy assessments with the child while the enumerator conducted the caregiver survey with the mother as well as questions that measure the children's leadership skills, noncognitive skills, and behavioral difficulties (see Appendix A, Section A-4 for details).

3.3 Outcomes

Our analysis focuses on the following groups of prespecified outcomes.

Learning outcomes. Children's learning outcomes were measured using an assessment test that comprises 15 questions on literacy (English & Bangla) and numeracy. All questions were taken from the national curriculum of Bangladesh. The test totals 80 points. We developed three sets of questions as we were covering students from different grades. The answers were set deliberately in binary form to avoid assessment bias. The questions of the assessment test are listed in Table A4.

Leadership, communication, and planning skills. We employed Scales for Rating the Behavioral Characteristics of Superior Students developed by Renzulli et al. (2002). This scale has 14 subscales. We took only leadership, communication, and planning subscales as our modules focus on these dimensions. Items of these three characteristics are listed in Table A5.

Noncognitive skills. We measured four types of noncognitive skills of the children. First, we measured the self-control of the children by using the Impulsivity Scale for Children (ISC), an 8-item survey that gives domain-specific students' impulsivity, defined as the "inability to regulate behavior, attention, and emotions in the service of valued goals" **(**Tsukayama, Duckworth, & Kim, 2013**)**. Second, we measured the grit of the participants using an 8-item grit scale developed by Duckworth and Quinn (2009). This scale measures perseverance – grit – as an individual difference score. Third, we measured the extent to which participating children view intelligence as a fixed behavioral trait rather than a feature that can be improved with effort using the 3-item growth mindset scale developed by Dweck, Chiu, and Hong (1995). Finally, we measured the impact of the intervention on the prosocial attitude of the children, using the Empathy Questionnaire for Children and Adolescents (EmQue-CA) developed by

Overgaauw, Rieffe, Broekhof, Crone, and Güroğlu (2017). Items of these scales are listed in Table A6.

Behavioral difficulties. We used the 25-item parent-report Strengths and Difficulties Questionnaire (SDQ) by Goodman (1997). The main motivation for using this scale was to check whether school closure triggered any fatigue, abnormality or conduct problem among the students. Furthermore, we wanted to investigate whether participation in this program reduced such problems by giving additional activities to do in household settings. Items of this scale are listed in Table A7.

Homeschooling time (student). Students' time investment in homeschooling.

Homeschooling time (caregiver): Caregiver involvement in children's educational activities.

All these outcome variables are standardized following Kling, Liebman, and Katz (2007). First, we normalized the raw values by subtracting the mean values of the Control group sample and then dividing by the standard deviation of the raw values in the Control group sample. All outcome variables are explained in the Appendix A, Section A-4.

3.4 Sample Descriptive Statistics

The average age of the children participating in this study was 7.4 years and the age range was 5.1 to 9.9 years. Around 51.3% of the children were girls. The mean years of schooling of their father and mother were 5.9 and 7.1 years, respectively. These children came from households that were mostly from a low socio-economic status with an average monthly income of BDT 11,003 (USD 130.8). Slightly more than half of the sample reported to have access to private tuition. In terms of access to distance learning modalities, these households were also quite disadvantaged. In May 2021, only 46.5% of them had access to TV, and less than 1% had access to computers or radio. All of them had access to mobile phones, 35.6% of which were smartphones. However, these smartphones were rarely used for educational purposes due to the high data costs and low internet speed, particularly in rural areas.

Table 1 and Table 2 present summary statistics and balance tests of various characteristics of our sample at the baseline and endline, respectively. In both cases, these characteristics are balanced across the treatment and control groups.

3.5 Usage

Providing access to distance education does not guarantee learning as students may not use the resources for many reasons, such as difficulties with navigating the system, technical

challenges with accessing the lessons, and time constraints of parents. It is therefore important to consider the level of student engagement with the audio lessons offered during our intervention.

Overall, participants' engagement was high: on average, each student listened to 7.3 hours of recorded lessons, which amounts to 43.6 minutes of listening per week. For reference, note that the audio content of this intervention was approximately 90 minutes per week. Around 70% of participating children completed more than two thirds of lessons in each of the modules (see Figure B3). Furthermore, there is a downward trend in the number of listeners as the program advances (see Figure B4). One potential reason is that some of the children might have lost interest after a few lessons or decided to skip some lessons for not finding them interesting or needed as they progressed. The other reason may be that schools were partially reopened in week 11 of our intervention, and this was announced a few weeks in advance. All in all, the number of lessons completed is moderately high, indicating that the intervention was well-received by the target group of households.

3.6 Empirical specification

To assess the overall effects of the treatments on the various outcomes, we estimate the following Ordinary Least Squares (OLS) regression specification:

$$Y_i = \alpha + \beta_1 T_{1i} + \beta_2 T_{2i} + \theta' X_i + \epsilon_i \qquad \text{(Equation 1)}$$

where Y_i is an outcome of a child from household i measured at the endline. The treatment indicator T_{1i} takes the value of one if the child is in the Standard group, and zero otherwise. The treatment indicator T_{2i} takes the value of one if the child is in the Extended group, and zero otherwise. The coefficients of interest are β_1 and β_2, which capture the causal effect of a treatment on an outcome. We also include a vector of individual and household-specific characteristics in the regression specification. X_i includes the child's age, gender, access to private tuition, parental education in years, family income, religion, access to TV, access to smartphone, homestead size, number of members in the household, and the relevant outcome measured at baseline. Finally, the error term ϵ_i captures all other unobserved influences.

We cluster the standard errors at the village level. We also separately report the Family Wise Error Rate (FWER) adjusted p-values corrected for multiple hypotheses testing using the free step-down resampling approach to account for the large set of outcomes that we considered in

this study (Westfall & Young, 1993). Furthermore, to account for uncertainty in the estimated treatment effects that arise naturally from the random assignment of participants into the treatments, we also report p-values using randomization-based inference (RI) (Young, 2019).

4 Results

4.1 Learning outcomes

Figure 4 summarizes the treatment effects of the intervention on the two main learning outcomes (literacy and numeracy), and their combined score (see Section A-4.1 for details on the assessments). Treatment effects are obtained from estimating Equation 1 with OLS. In panel A, which considers the full sample, we see that treated children in the Standard and Extended groups experienced a 0.60 standard deviation (SD) and 0.63 SD improvement in total score relative to children in the control group, respectively. In the literacy component, treatment effects were 0.55 SD ($p<0.01$) and 0.59 SD ($p<0.01$) for the Standard and Extended treatment arms, respectively. In the numeracy component, treatment effects were 0.53 SD ($p<0.01$) and 0.54 SD ($p<0.01$) for the Standard and Extended treatment arms, respectively (see Table 3 and Table 4 for details). In terms of raw test scores, the treatment groups achieved about 30% higher total scores than the control group, which translates to about 2.5 additional correct answers (out of 15), when the average number of correct answers in the control group was 8 (see Table B3).

These results indicate that the provision of audio lessons in a context where no access to formal education is available can result in significant improvements in students' learning outcomes in both literacy and numeracy (see Figure B5 for comparisons of the distribution of test scores across treatment groups).

We also estimate the dosage-response relationship between the number of IAI lessons accessed and the main outcomes of interest (test scores). We first estimate the relationship using OLS. Because the number of lessons was decided by the caregiver/child after the randomization, we also estimate the dosage-response relationship using an instrumental variable approach where treatment assignment serves as the instrument. These results are presented in Table 5. In column 1, we present the OLS estimates where the key explanatory variable is the number of lessons. We find that each lesson increases test score by 0.01 SD. In column 2, we present instrumental variable estimates where the number of lessons is instrumented with a treatment dummy. We find that each lesson increases total test score by 0.02 SD, literacy score by 0.03

SD, and numeracy score by 0.03 SD. Given that the average number of lessons taken among the treated groups is 20.42 for literacy and 20.13 for numeracy, these effects translate to 0.58 SD for literacy score and 0.54 SD for numeracy score.

Panels B and C of Figure 4 show the treatment effects by gender. The intervention seems to have benefitted equally boys' and girls' test scores. This is confirmed through regression analysis in which treatment is interacted with gender (see Table 8).

4.2 Leadership

The Extended treatment group was offered an additional module that focused on leadership, qualities of a leader, active listening, communication and presenting, planning, and so on. At the endline, we collected measures of these skills using scales for Rating the Behavioral Characteristics of Superior Students (see Section A-4.2) (Renzulli et al., 2002). Treatment effects on these measures are presented in Figure 5, Panel A. We do not find any evidence that children in the Extended treatment improved on these measures relative to the control group—treatment effects are small and statistically insignificant. This suggests that improving leadership skills might be difficult to achieve through this distance learning medium for children of this age.

4.3 Noncognitive skills

While learning outcomes are the core targets of this educational intervention, the development of noncognitive skills, such as critical thinking, problem-solving skills, social skills, persistence, creativity, self-control, etc. is an integral part of any educational program. Therefore, in our intervention, through various examples and discussions, we tried to develop the noncognitive skills of treated children. During the endline, we measured the level of noncognitive skills of the children. We chose commonly used noncognitive skill measurement scales, e.g., impulsivity, grit, growth mindset, and empathy instruments (see Section A-4.3).

Treatment effects on noncognitive skills are presented in Figure 5, Panel B. We find that our treatments were effective in reducing the impulsive behavior of children in the treatment groups. In the overall impulsivity component, we estimate treatment effects of -0.27 SD ($p<0.05$) and -0.15 SD ($p<0.05$) for the Standard and Extended treatments, respectively. However, these effects are not robust when we consider FWER adjusted p-values (Table 4). We also do not find any significant treatment effects on the other three measures: grit, growth mindset and level of empathy ($p>0.05$).

4.4 Behavioral Difficulties

Our intervention started at a time when the children were not attending school, meaning that they had not been following a formal educational routine for about 15 months. Because of the lack of school attachment, one might be concerned that children could develop behavioral problems, e.g., tantrums, nervousness, lack of control over emotions, and hyperactivity. These problems might be reduced by participating in our intervention. To check this possibility, we use the 25-item parent-report Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). This popular scale covers five domains of children's behaviors, i.e., emotional symptoms, conduct problems, hyperactivity, peer problem and prosociality (see Section A-4.4).

In all the five domains, we find statistically significant treatment effects for children in the Standard treatment (Figure 5, Panel C). Children in this group showed less emotional symptoms, conduct problems, hyperactivity, peer problems and a more prosocial attitude compared to the children in the control group. In the Standard group, treatment effects in various domains of SDQ range from -0.17 SD to -0.24 SD ($p<0.05$). However, these effects are not robust when we consider FWER adjusted p-values (Table 4). On the other hand, in the Extended group, treatment effects in various domains of SDQ range from -0.02 SD to -0.13 SD but are not statistically significant whether we consider p-values based on robust standard errors clustered at the village level, FWER adjusted p-values, or RI p-values. One potential explanation for the noisier estimated effects is that the greater number of activities provided in the Extended group might unintentionally disrupt the flow of learning and reduce any potential associated benefits.

Regular extended interaction between caregiver and child may be one of the potential underlying drivers of some of the effects on behavioral issues. As will be shown below, in our intervention caregivers not only interacted with their child during the lesson play but also engaged in some other activities with them during the rest of the day. These findings suggest that even if distance education is minimal and low-tech, by establishing a routine and introducing engaging activities to the rural children's life it might eventually help them avoid behavioral difficulties.

4.5 Heterogeneous effects

In Figure 6, we present heterogeneous effects of the intervention along various dimensions. The top subfigure presents the mean test score differences between the two treatment groups

and the control group by quartile of baseline test scores. Differences in test scores are the largest for those who had scored the lowest in the baseline. For students in the 1st and 2nd quartiles (weakest), the gaps between treatment and control groups are as large as 40%, whereas, in the other two quartiles, the gaps are less than 30%. All these gaps are statistically significant at the 5% level. These results suggest that the program particularly helped academically poorer students to catch up on their learning gaps.

The middle subfigure of Figure 6 shows the total score by family income. Again, the treatment effects are the largest for children who come from households in lower-income quartiles. The gains in test scores are more than 35% in the 1st and 2nd quartiles, whereas they are around 20% in the top two quartiles. This indicates that the largest learning gains occur for the poorest strata of the sample, as children in these groups might have been struggling the most due to the lack of access to alternative schooling during the pandemic-induced school closures.

The bottom subfigure of Figure 6 shows the heterogeneity in the treatment effects based on parental education. We find that treatment effects are significantly higher for students with less-educated parents in both treatment groups.

Overall, these results suggest that our intervention was especially beneficial for academically weaker students, students from the poorest strata, and students with less-educated caregivers (see also Table 8).

4.6 Robustness

Social Desirability Bias

As some of our outcome variables were collected via surveys, social desirability bias may arise, which can lead to over or underestimation of the treatment effects on these outcomes. To address this concern, during the endline, we elicited the social desirability bias of parents using the short-form Crowne-Marlowe module (Crowne & Marlowe, 1960; Reynolds, 1982) and of the children using the Children's Social Desirability (CSD-S) scale (Miller et al., 2015) (see Section A-5). In the Appendix B, Table B4 and Table B5 we report analysis that suggests that the significant effects of the intervention are mainly driven by participants with lower social desirability bias, thereby alleviating concerns about social desirability bias driving our results.

Assessment Reliability

Our assessment tests are based on the curriculum designed by the National Curriculum and Textbook Board (NCTB) of Bangladesh. This curriculum has been already rigorously tested

and modified over the years based on the changing aptitude of the learners. We picked 15 questions from the latest version of the textbooks for primary grades. To verify the coherence of these selected items, we report some statistical tests (see Section A-5) that indicate that our assessment tests were stable, reliable, and coherent.

5 Channels: Did the Intervention Impact Homeschooling Time Investment?

We next examine the impact of the intervention on time spent homeschooling, which we split into two components: students' study time and caregivers' homeschooling time. These variables are defined in the Appendix A, Section A-4.5 and A-4.6.

We do not find a statistically significant treatment effect on the student's study time (See Table 6 and Table 7). Our sample children were in primary school during the intervention. Self-study is not very common for children of these ages, which probably explains why we do not find a significant impact of the intervention on students' study time. Instead, we find that caregivers who participated in the intervention devoted more time to their children's education-related tasks compared to caregivers in the Control group — 10.4 minutes per day in the Standard group and 3.1 minutes per day in the Extended group. Treatment effects in the Standard and Extended groups were 0.20 SD ($p<0.01$) and 0.07 ($p<0.05$), respectively. The effects measured in Likert scale are generally more robust than the effects measured in minutes when we consider FWER adjusted p-values (Table 7). These results suggest that the IVR-based education program encouraged caregivers to engage more in their children's education, even outside of the direct program time.

6 Discussion

The intervention was highly effective in improving the learning outcomes of children in the treatment groups, especially those from a low-socioeconomic background. The effect sizes are within the range of other educational interventions implemented during the Covid-19 pandemic. On the one hand, they are somewhat larger than the effect of an eight-week after-school online math tutoring program provided to secondary school students in highly disadvantaged neighborhoods in Spain six months after Covid-19 school closures ended (Gortazar, Hupkau, & Roldán, 2022), a five-week volunteer-based online tutoring in Italy (Carlana & La Ferrara, 2021), and a 12-week pilot program of online tutoring by college-volunteers in the US (Kraft, List, Livingston, & Sadoff, 2022). The effect sizes are also larger

than the effect of an eight-week direct phone call based tutoring in Botswana (Angrist et al., 2022). On the other hand, they are a bit smaller than the effect of a 13-week phone-based mentoring program provided to primary school age children in rural Bangladesh by volunteers during Covid-19 school closures (Hassan et al., 2021).

The large effects found here are likely due to the 18-month-long school closures in Bangladesh that prevented children in the control group from having access to any educational services. One might expect that the large differences between the treatment and control groups would diminish over time after children in the control group started having access to formal education. For instance, it has been reported that in a similar setting, one year after a phone-based tutoring intervention ended and when children in the control group started attending school, the effect decreased by roughly 20 to 55 percent (Hassan et al., 2021).

There is weak evidence that the intervention also reduced impulsivity and behavioral difficulties for children in the Standard treatment. These findings suggest that remote learning can also potentially be used to address some behavioral issues of children, particularly when schools are closed, and children have limited activities to do at home. Specifically, the intervention required children to go through the module activities and interact with their caregivers. Going through this process on a regular basis might then help to reduce impulsivity, emotional symptoms, conduct problems, and hyperactivity, as well as improve prosocial behavior and ability to interact with others.

Based on the exchange rate of 80 BDT to 1 USD (at the time of the intervention), the 15-week intervention cost USD 27.5 per student, of which USD 13.2 was the variable cost and the remaining USD 14.3 was the fixed cost. Thus, the intervention is among one of the most cost-effective interventions implemented during COVID-19 school closures (see, for example, Hassan et al., 2021 and Angrist et al., 2022, for cost effectiveness of those interventions). If the duration of the intervention is lengthened, the total cost would increase proportionately to our 15-week intervention. As only a total of 1,182 students in two districts received this intervention, the fixed cost per student was high relative to the variable cost. Scaling up the intervention to more students will lower the per-student fixed cost. It is also possible to lower the variable cost if a lower phone call rate can be negotiated with telecommunication companies. Note that these cost figures include both provider's and recipients' cost.

7 Conclusion

The household environment plays a critical role in education, but most education policies primarily focus on school-based interventions, as it is believed that it is more feasible to improve schools than to intervene at the household level at scale (Muralidharan & Singh, 2021). The school closures induced by the Covid-19 pandemic has sharply shifted the focus from the school to the household environment. Due to the weak information communication technology's ecosystem in most low-income developing countries, widely accessible basic feature phones have become popular in educating the mass of students during the Covid-19 pandemic (Hassan et al., 2021). Existing studies using basic feature phones are limited to SMS reminders or brief calls to the parents to follow up on their children's homework (Angrist et al., 2022; Lichand & Christen, 2021; Muralidharan & Singh, 2021).

In this paper, using the IVR system, we offer an extension of the existing applications of basic feature phones in education. We delivered these lessons via basic mobile phones because the basic mobile phone penetration rate in rural Bangladesh is significantly higher than other one-way technologies such as radio and television. Our results indicate that this approach delivered substantial learning benefits to students, especially for those who are more disadvantaged.

Although the extent of learning disruptions of the Covid-19 pandemic were unprecedented, educational disruptions at a smaller scale are not uncommon in low- and middle-income countries. In many developing countries, climate change, natural and human-induced events (e.g., cyclones, floods, wars, and political unrest) often damage educational infrastructure and limit school operations. Thus, policymakers may consider expanding education delivery in out-of-school settings using accessible distant learning methods, such as the one studied in this paper, to better support children's learning in these situations. Importantly, the relatively low cost of the intervention examined in this study, especially if it is provided to large numbers of children, makes it a promising option for providing remedial educational support to poor and academically left-behind students in hard-to-reach areas, even outside of times of crisis. A fruitful avenue for future research would be to examine the effectiveness of this type of intervention in such settings.

References

Agostinelli, F., Doepke, M., Sorrenti, G., & Zilibotti, F. (2022). When the great equalizer shuts down: Schools, peers, and parents in pandemic times. *Journal of public economics, 206*, 104574.

Aker, J. C., & Ksoll, C. (2019). Call me educated: Evidence from a mobile phone experiment in Niger☆. *Economics of Education Review, 72*, 239-257.

Angrist, N., Bergman, P., & Matsheng, M. (2022). Experimental evidence on learning using low-tech when school is out. *Nature Human Behaviour, 6*(7), 941-950.

Anzalone, S., & Bosch, A. (2005). Improving educational quality with interactive radio instruction: a toolkit for policymakers and planners. *Working Paper*.

Banerjee, A., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The quarterly journal of economics, 122*(3), 1235-1264.

BBS. (2017). Household Income and Expenditure Survey 2016–2017. In. Dhaka, Bangladesh: Ministry of Planning, Government of the People's Republic of Bangladesh.

Beam, E. A., Mukherjee, P., Navarro-Sola, L., Ferdosh, J., & Sarwar, M. A. H. (2021). Take-Up, Use, and Effectiveness of Remote Technologies. In *Working Paper Series*. Washington, DC: The World Bank.

Bergman, P. (2021). Parent-child information frictions and human capital investment: Evidence from a field experiment. *Journal of political economy, 129*(1), 000-000.

Berlinski, S., Busso, M., Dinkelman, T., & Martinez, C. (2016). Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile. *Unpublished manuscript*.

Bird, K. A., Castleman, B. L., Denning, J. T., Goodman, J., Lamberton, C., & Rosinger, K. O. (2021). Nudging at scale: Experimental evidence from FAFSA completion campaigns. *Journal of Economic Behavior & Organization, 183*, 105-128.

Biswas, K., Asaduzzaman, T., Evans, D. K., Fehrler, S., Ramachandran, D., & Sabarwal, S. (2020). TV-Based Learning in Bangladesh: Is it Reaching Students? In *Working Paper Series*. Washington, DC: The World Bank.

Bosch, A. (1997). Interactive Radio Instruction: Twenty-three years of improving education quality. In *Washington, DC: World Bank Group*.

Bosch, A., Rhodes, R., & Kariuki, S. (2002). Interactive Radio Instruction: an Update from the Field. In W. D. Haddad & A. Draxler (Eds.), *Technologies for Education* (pp. 134): Knowledge Enterprise, Inc.

Carlana, M., & La Ferrara, E. (2021). Apart but Connected: Online Tutoring and Student Outcomes during the COVID-19 Pandemic. In *HKS Faculty Research Working Paper Series*

Castleman, B. L., & Meyer, K. E. (2020). Can text message nudges improve academic outcomes in college? Evidence from a West Virginia initiative. *The Review of Higher Education, 43*(4), 1125-1165.

Crawfurd, L., Evans, D. K., Hares, S., & Sandefur, J. (2021). Teaching and testing by phone in a pandemic. In *CGD Working Paper*. Washington, DC: Center for Global Development.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of consulting psychology, 24*(4), 349.

DataReportal. (2021). Digital 2021: Bangladesh. Retrieved from https://datareportal.com/reports/digital-2021-bangladesh

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of personality assessment, 91*(2), 166-174.

Dweck, C. S., Chiu, C.-y., & Hong, Y.-y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological inquiry, 6*(4), 267-285.

Elliott, S. N., Kratochwill, T. R. L., & Travers, J. F. (2000). *Educational psychology: Effective teaching effective learning*. Boston: McGraw-Hill.

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences, 118*(17), e2022376118.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry, 38*(5), 581-586.
Gortazar, L., Hupkau, C., & Roldán, A. (2022). Online tutoring works: Experimental evidence from a program with vulnerable children. *Working Paper*.
Hassan, H., Islam, A., Siddique, A., & Wang, L. C. (2021). Telementoring and homeschooling during school closures: A randomized experiment in rural Bangladesh. *Working Paper*.
Ho, J., & Thukral, H. (2009). Tuned in to student success: Assessing the impact of interactive radio instruction for the hardest-to-reach. *Journal of Education for International Development, 4*(2), 34-51.
Hurwitz, L. B., Lauricella, A. R., Hanson, A., Raden, A., & Wartella, E. (2015). Supporting Head Start parents: impact of a text message intervention on parent–child activity engagement. *Early Child Development and Care, 185*(9), 1373-1389.
Islam, A., Malek, A., Tasneem, S., & Wang, L. C. (2022). Can public recognition reward backfire? Field experimental evidence on the retention and performance of volunteers.
Karagianni, D., & Jude Montgomery, A. (2018). Developing leadership skills among adolescents and young adults: a review of leadership programmes. *International Journal of Adolescence and Youth, 23*(1), 86-98.
Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica, 75*(1), 83-119.
Kraft, M. A., & Dougherty, S. M. (2013). The effect of teacher–family communication on student engagement: Evidence from a randomized field experiment. *Journal of Research on Educational Effectiveness, 6*(3), 199-222.
Kraft, M. A., List, J. A., Livingston, J. A., & Sadoff, S. (2022). Online Tutoring by College Volunteers: Experimental Evidence from a Pilot Program. *AEA Papers and Proceedings, 112*, 614-618. doi:10.1257/pandp.20221038
Ksoll, C., Aker, J. C., Miller, D., Perez, K., & Smalley, S. L. (2015). Learning without Teachers? Evidence from a Randomized Experiment of a Mobile Phone-Based Adult Education Program in Los Angeles. *Working Paper*.
LEAD. (2021). LEAD Learning. *2021*(15 January). Retrieved from https://leadedu.org/learning
Lichand, G., & Christen, J. (2021). Behavioral nudges prevent student dropouts in the pandemic. In *ECON - Working Papers, Department of Economics - University of Zurich*.
Mayer, S. E., Kalil, A., Oreopoulos, P., & Gallegos, S. (2015). *Using behavioral insights to increase parental engagement: The parents and children together (PACT) intervention*. Retrieved from
Miller, P. H., Baxter, S. D., Royer, J. A., Hitchcock, D. B., Smith, A. F., Collins, K. L., . . . Vaadi, K. K. (2015). Children's social desirability: Effects of test assessment mode. *Personality and individual differences, 83*, 85-90.
Moscoviz, L., & Evans, D. K. (2022). Learning loss and student dropouts during the covid-19 pandemic: A review of the evidence two years after schools shut down. In *Center for Global Development, Working Paper* (Vol. 609).
Muralidharan, K., & Singh, A. (2021). India's new National Education Policy: Evidence and challenges. *Science, 372*(6537), 36-38. doi:10.1126/science.abf6655
Overgaauw, S., Rieffe, C., Broekhof, E., Crone, E. A., & Güroğlu, B. (2017). Assessing empathy across childhood and adolescence: Validation of the Empathy Questionnaire for Children and Adolescents (EmQue-CA). *Frontiers in psychology, 8*, 870.
Patrinos, H. A. (2022). Learning loss and learning recovery. *DECISION*, 1-6.
Rahman, H. Z., Matin, I., Rahman, A., Das, N. C., Zillur, U., Ahmed, M. S., . . . Shameem, N. (2021). PPRC-BIGD COVID-19 Livelihoods & Recovery Panel Survey. In (Vol. 2021). Dhaka, Bangladesh: BIGD.
Rahman, T., & Sharma, U. (2021). A Simulation of COVID-19 School Closure Impact on Student Learning in Bangladesh. In. Washington, DC: The World Bank.
Renzulli, J. S., Smith, L. H., White, A. J., Callahan, C. M., Hartman, R. K., & Westberg, K. L. (2002). *Scales for rating the behavioral characteristics of superior students. Technical and administration manual*: ERIC.
Reynolds, W. M. (1982). Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *Journal of clinical psychology, 38*(1), 119-125.

Schueler, B. E., & Rodriguez-Segura, D. (2022). A Cautionary Tale of Tutoring Hard-to-Reach Students in Kenya. *Journal of Research on Educational Effectiveness*.

Singh, A., Romero, M., & Muralidharan, K. (2022). COVID-19 Learning Loss and Recovery: Panel Data Evidence from India. In: National Bureau of Economic Research.

Sisk, D. A. (1993). Leadership education for the gifted. In K. A. Heller, F. J. Mönks, & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 491–505). New York, NY: Pergamon.

Tsukayama, E., Duckworth, A. L., & Kim, B. (2013). Domain-specific impulsivity in school-age children. *Developmental Science, 16*(6), 879-893.

UNESCO. (2022). *UNESCO map on school closures and UIS*. Retrieved from: https://en.unesco.org/covid19/educationresponse and http://data.uis.unesco.org

UNICEF. (2020). Students in Bangladesh adjust to remote learning via national TV during COVID-19 lockdown. Retrieved from https://www.unicef.org/bangladesh/en/stories/students-bangladesh-adjust-remote-learning-national-tv-during-covid-19-lockdown

Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment* (Vol. 279): John Wiley & Sons.

Young, A. (2019). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The quarterly journal of economics, 134*(2), 557-598.

Figures and Tables

Figure 1. Summary of Research Design.

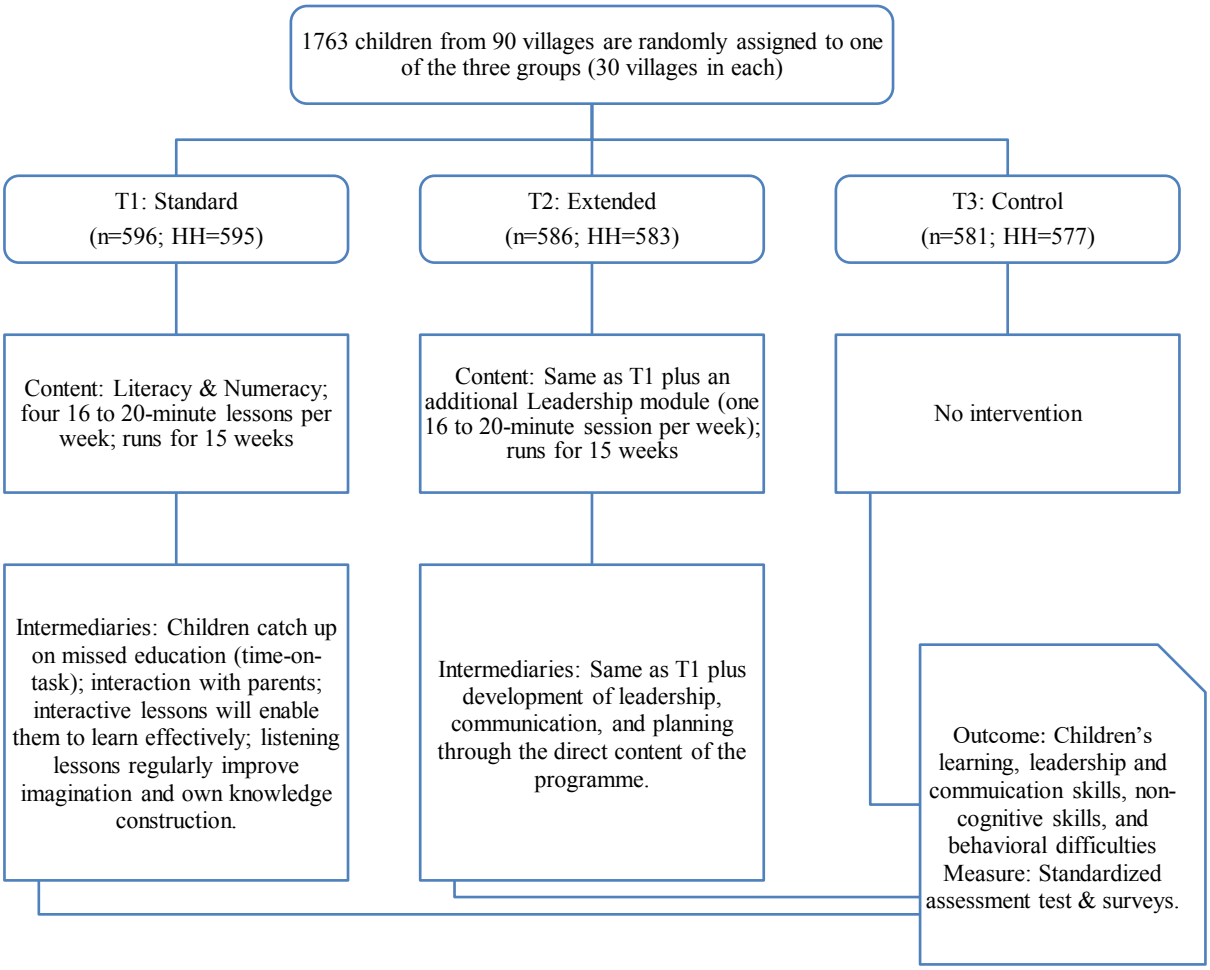
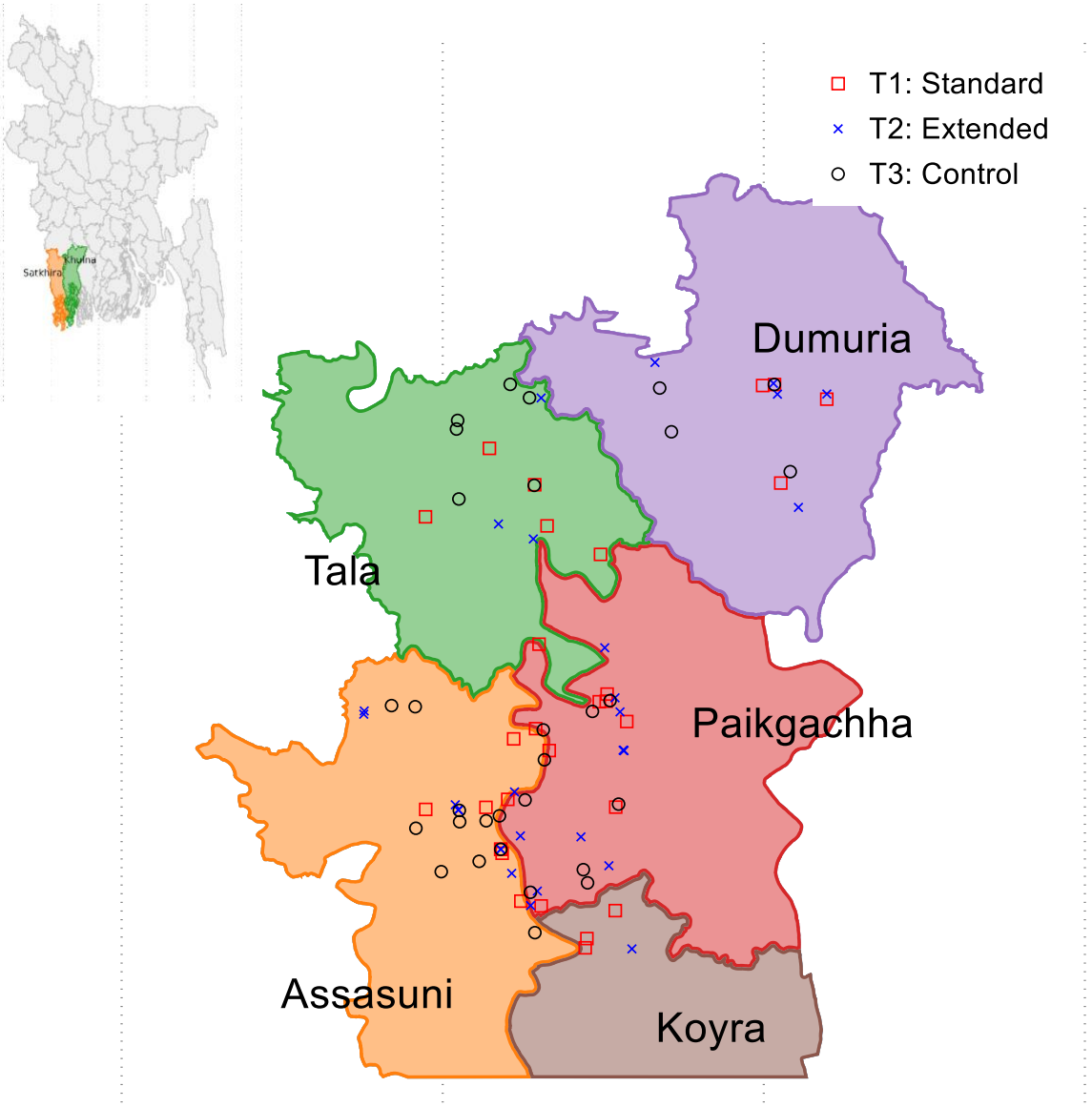


Figure 2. Bangladesh map, study area and treatment villages.

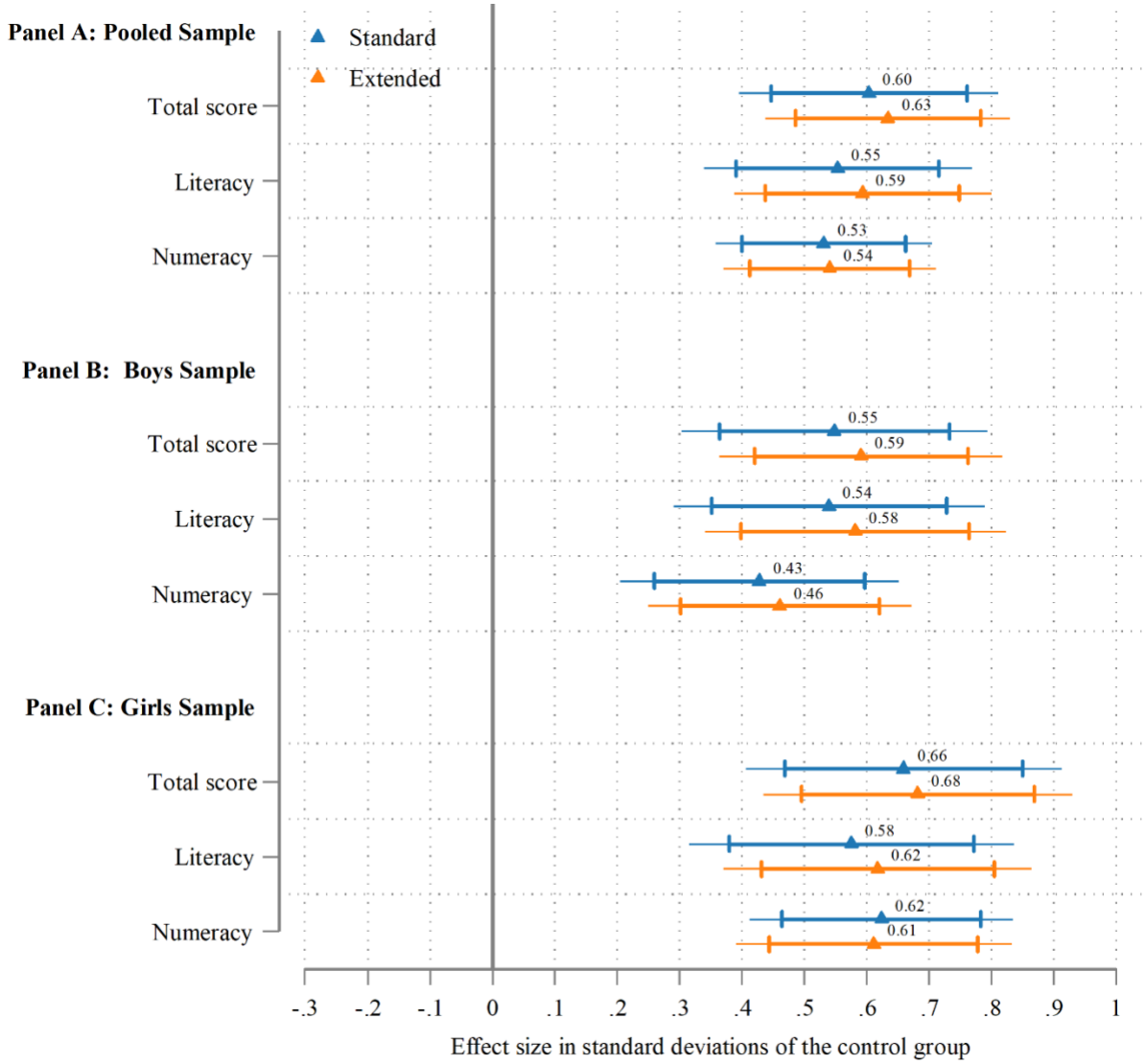


Note: This figure exhibits the study district and subdistricts of Bangladesh. Markers indicate the villages (clusters) of this intervention.

Figure 3. Project Timeline.

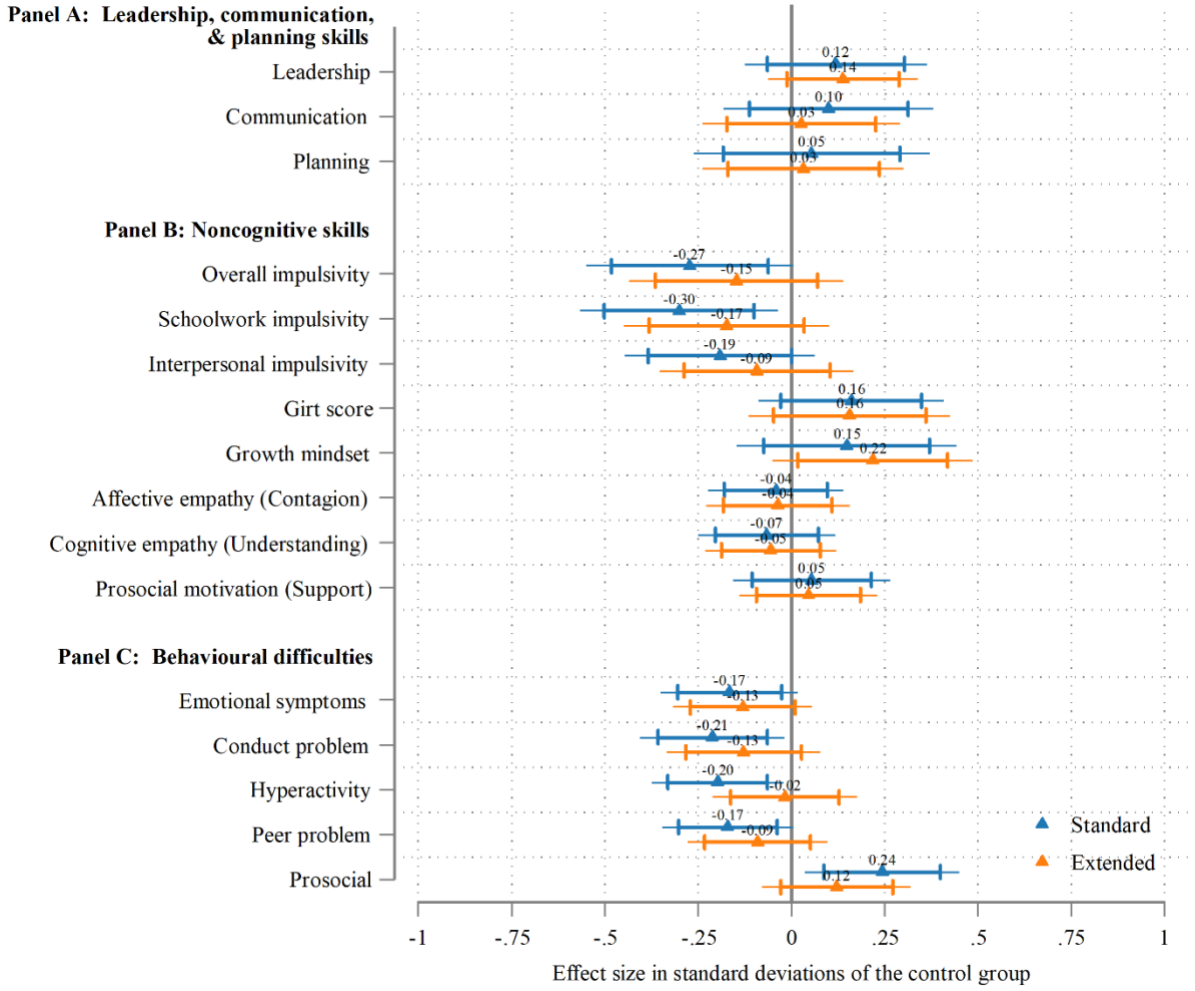
Preparation ↓	Baseline ↓	Random assignment ↓	Intervention start ↓	Intervention end ↓	Endline ↓
Mar to May 21	*May 21*	*Jun 21*	*Jun 21*	*Oct 21*	*Nov 21*
Content design; lesson recordings; IVR flow design & optimization; field test.	Baseline survey	Randomization; mothers' briefing	15-week intervention period		Children's assessment test; children's survey; parent's survey

Figure 4. Treatment effects on the standardized learning outcomes.



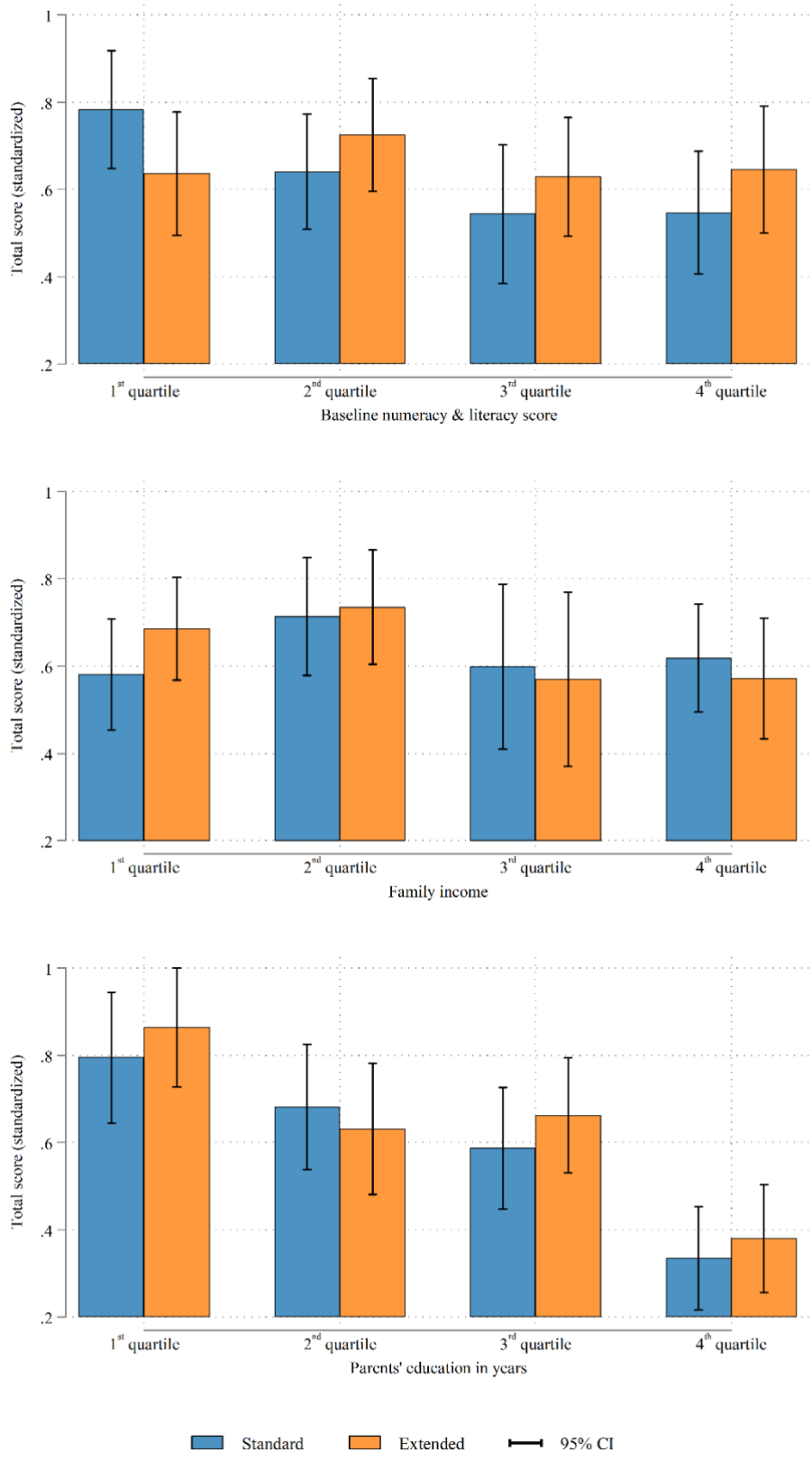
Note: All outcomes are standardized indices with the control group having a mean of 0 and SD of 1. Therefore, this figure shows where the mean of the treatment groups lies in the distribution of the control group in standard deviation (SD) units, with 95 and 99 confidence intervals. Coefficients are estimated using OLS. Baseline controls included: children's age, baseline literacy score, baseline numeracy score, access to private tuition, parents' education in years, family income, religion, access to TV & smartphone, homestead size and the number of members in the household. Children's grade fixed effects are included in all regressions. Robust standard errors are clustered at the village level.

Figure 5. Treatment effects on secondary outcomes



Note: All outcomes are standardized indices with the control group having a mean of 0 and SD of 1. Therefore, this figure shows where the mean of the treatment groups lies in the distribution of the control group in standard deviation (SD) units, with 95 and 99 confidence intervals. Coefficients are estimated using OLS. Baseline controls included: children's age, baseline literacy score, baseline numeracy score, access to private tuition, parents' education in years, family income, religion, access to TV & smartphone, homestead size and the number of members in the household. Children's grade fixed effects are included in all regressions. Robust standard errors are clustered at the village level.

Figure 6. Differences in assessment test performance, by baseline numeracy & literacy score, family income, parents' education quartiles.



Note: This figure exhibits the difference in assessment test performance between treatment and control groups, by quartiles of baseline numeracy & literacy score, family income, and parents' education for all children who completed the endline assessment tests.

Table 1. Sample characteristics and balance at baseline.

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
	T1: Standard	T2: Extended	T3: Control	Full sample	*P-Value*
Children's age in years (as of 01/01/2021)	7.38 (0.03)	7.39 (0.03)	7.38 (0.03)	7.38 (0.02)	0.90
Gender (=1 if Boy)	0.49 (0.02)	0.49 (0.02)	0.48 (0.02)	0.48 (0.01)	0.74
Baseline literacy score	16.98 (0.32)	16.63 (0.23)	16.86 (0.35)	16.82 (0.10)	0.89
Baseline numeracy score	14.86 (0.20)	14.66 (0.17)	14.84 (0.17)	14.79 (0.07)	0.70
Access to private tuition (=1 if yes)	0.51 (0.03)	0.58 (0.04)	0.56 (0.04)	0.55 (0.01)	0.71
Father's education (in years)	6.25 (0.25)	5.85 (0.22)	5.75 (0.24)	5.95 (0.10)	0.32
Mother's education (in years)	7.25 (0.24)	6.96 (0.18)	7.01 (0.19)	7.07 (0.08)	0.69
Household member	4.82 (0.12)	4.76 (0.07)	4.89 (0.08)	4.83 (0.04)	0.34
Family income (in BDT/ month)	10977.19 (295.24)	10900.26 (349.99)	11189.67 (433.44)	11021.64 (129.70)	0.61
Access to TV (=1 if yes)	0.46 (0.03)	0.45 (0.04)	0.48 (0.03)	0.46 (0.01)	0.48
Access to smartphone (=1 if yes)	0.34 (0.03)	0.36 (0.02)	0.36 (0.03)	0.36 (0.01)	0.73
Homestead land (in decimal)	9.14 (0.60)	9.56 (0.86)	11.26 (1.11)	9.98 (0.43)	0.12
Religion (=1 if Islam)	0.77 (0.05)	0.81 (0.04)	0.74 (0.05)	0.77 (0.01)	0.48
Observation (Households)	596 (595)	586 (583)	581 (577)	1763 (1755)	-
Village	30	30	30	90	-

Note: This table reports the background characteristics of the participants of different treatment groups and for all participants during the baseline. The rightmost column reports the p-value from the F-test of joint significance. Robust standard errors clustered at the village level are in parentheses.

Table 2. Sample characteristics and balance at endline.

Variables	(1) T1: Standard	(2) T2: Extended	(3) T3: Control	(4) Full sample	(5) *P-Value*
Children's age in years (as of 01/01/2021)	7.39 (0.04)	7.38 (0.03)	7.38 (0.04)	7.39 (0.02)	0.98
Gender (=1 if Boy)	0.49 (0.02)	0.48 (0.02)	0.49 (0.02)	0.49 (0.01)	0.96
Baseline literacy score	17.03 (0.33)	16.68 (0.25)	16.87 (0.35)	16.86 (0.10)	0.97
Baseline numeracy score	14.93 (0.21)	14.70 (0.17)	14.85 (0.17)	14.83 (0.07)	0.87
Access to private tuition (=1 if yes)	0.51 (0.03)	0.59 (0.04)	0.57 (0.04)	0.55 (0.01)	0.69
Father's education (in years)	6.21 (0.27)	5.84 (0.22)	5.72 (0.24)	5.93 (0.10)	0.32
Mother's education (in years)	7.25 (0.24)	6.95 (0.18)	7.00 (0.19)	7.07 (0.08)	0.67
Household member	4.82 (0.13)	4.75 (0.07)	4.91 (0.08)	4.83 (0.04)	0.25
Family income (in BDT/ month)	10963.24 (297.76)	10852.40 (347.97)	11196.26 (438.98)	11003.73 (133.60)	0.56
Access to TV (=1 if yes)	0.46 (0.03)	0.46 (0.04)	0.48 (0.03)	0.47 (0.01)	0.54
Access to smartphone (=1 if yes)	0.34 (0.03)	0.36 (0.02)	0.35 (0.03)	0.35 (0.01)	0.93
Homestead land (in decimal)	9.13 (0.61)	9.41 (0.81)	11.12 (1.12)	9.89 (0.44)	0.14
Religion (=1 if Islam)	0.80 (0.05)	0.83 (0.04)	0.77 (0.05)	0.80 (0.01)	0.45
Observation (Households)	567 (566)	562 (560)	561 (558)	1690 (1684)	-
Village	30	30	30	90	-

Note: This table reports the background characteristics of the participants of different groups and for all participants who participated in the endline survey and assessment. The rightmost column reports the p-value from the F-test of joint significance. Robust standard errors clustered at the village level are in parentheses.

Table 3. Means and differences of key outcomes, by treatment.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	T1: Standard	T2: Extended	T3: Control	Difference (T1 & T3)	Difference (T2 & T3)	Difference (T1 & T2)	Hypothe-sized sign
Total score	56.11	56.38	42.75	13.36***	13.63***	-0.27	(+)
	(1.23)	(1.03)	(1.63)	(2.05)	(1.94)	(1.61)	
				[31.26%]	[31.88%]	[-0.47%]	
Literacy	33.50	33.79	25.29	8.21***	8.50***	-0.29	(+)
	(0.90)	(0.74)	(1.12)	(1.44)	(1.35)	(1.17)	
				[32.48%]	[33.63%]	[-0.86%]	
Numeracy	22.61	22.59	17.47	5.15***	5.13***	0.02	(+)
	(0.42)	(0.44)	(0.58)	(0.72)	(0.73)	(0.61)	
				[29.48%]	[29.35%]	[0.10%]	
Overall impulsivity	1.90	1.99	2.10	-0.21**	-0.12	-0.09	(-)
	(0.06)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	
				[-9.85%]	[-5.57%]	[-4.53%]	
Schoolwork impulsivity	1.97	2.07	2.22	-0.25***	-0.15	-0.10	(-)
	(0.06)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	
				[-11.36%]	[-6.69%]	[-5.01%]	
Interpersonal impulsivity	1.82	1.90	1.99	-0.16*	-0.09	-0.08	(-)
	(0.06)	(0.07)	(0.06)	(0.09)	(0.09)	(0.09)	
				[-8.16%]	[-4.32%]	[-4.01%]	
Grit score	3.15	3.14	3.05	0.10*	0.09	0.00	(+)
	(0.04)	(0.05)	(0.04)	(0.06)	(0.06)	(0.06)	
				[3.19%]	[3.06%]	[0.12%]	
Growth mindset	3.69	3.78	3.51	0.18	0.28**	-0.09	(+)
	(0.10)	(0.08)	(0.11)	(0.14)	(0.13)	(0.13)	
				[5.25%]	[7.94%]	[-2.50%]	
Affective empathy (contagion)	1.14	1.14	1.15	-0.02	-0.01	-0.00	(+)
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	
				[-1.36%]	[-1.06%]	[-0.31%]	
Cognitive empathy (understanding)	0.85	0.85	0.88	-0.03	-0.03	-0.00	(+)
	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	
				[-3.30%]	[-3.21%]	[-0.10%]	
Prosocial motivation (support)	1.32	1.31	1.29	0.02	0.02	0.01	(+)
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	
				[1.85%]	[1.18%]	[0.67%]	
Leadership	28.72	28.82	27.99	0.72	0.83*	-0.11	(+)
	(0.46)	(0.31)	(0.36)	(0.59)	(0.48)	(0.55)	
				[2.59%]	[2.97%]	[-0.37%]	
Communication	15.27	14.96	14.82	0.45	0.13	0.31	(+)
	(0.35)	(0.30)	(0.31)	(0.47)	(0.43)	(0.46)	
				[3.02%]	[0.90%]	[2.10%]	
Planning	50.85	50.48	50.04	0.81	0.44	0.37	(+)
	(1.36)	(1.00)	(1.05)	(1.72)	(1.45)	(1.69)	
				[1.62%]	[0.89%]	[0.73%]	
Emotional symptoms	2.40	2.50	2.73	-0.33**	-0.23	-0.10	(-)
	(0.09)	(0.09)	(0.11)	(0.14)	(0.14)	(0.13)	
				[-11.93%]	[-8.34%]	[-3.92%]	
Conduct problem	2.24	2.40	2.67	-0.43***	-0.27*	-0.16	(-)
	(0.09)	(0.10)	(0.12)	(0.15)	(0.16)	(0.13)	
				[-15.93%]	[-9.98%]	[-6.61%]	
Hyperactivity	3.98	4.36	4.40	-0.42***	-0.04	-0.38**	(-)
	(0.10)	(0.13)	(0.11)	(0.15)	(0.17)	(0.16)	
				[-9.52%]	[-0.87%]	[-8.73%]	

Peer problem	2.54 (0.07)	2.66 (0.08)	2.78 (0.07)	-0.24** (0.10) [-8.66%]	-0.13 (0.10) [-4.52%]	-0.12 (0.10) [-4.33%]	(-)
Prosocial	7.55 (0.11)	7.31 (0.10)	7.06 (0.12)	0.49*** (0.17) [6.96%]	0.24 (0.16) [3.45%]	0.25 (0.15) [3.39%]	(+)
Observation	567	562	561	1128	1123	1129	-
Village	30	30	30	60	60	60	-

Note: In columns 1-3, the means of each of the three treatment arms are presented. Columns 4-6 present estimates of univariate OLS regressions, where the omitted treatment category is T3 in column 4, T3 in column 5, and T2 in column 6. Column 7 exhibits the hypothesized direction of the differences in columns 4 and 5. There was no a priori assumption about the difference between the two treatment arms. Robust standard errors clustered at the village level are in parentheses. Percentages of change between groups marked in column-header are in square brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Treatment effect on key outcomes.

Variables	T1: Standard				T2: Extended			
	(1) Coefficient on treatment dummy	(2) t-test *P-value*	(3) FWER *P-value*	(4) RI *P-value*	(5) Coefficient on treatment dummy	(6) t-test *P-value*	(7) FWER *P-value*	(8) RI *P-value*
Total score	0.60	0.00	0.00	0.00	0.63	0.00	0.00	0.00
Literacy	0.55	0.00	0.00	0.00	0.59	0.00	0.00	0.00
Numeracy	0.53	0.00	0.00	0.00	0.54	0.00	0.00	0.00
Overall impulsivity	-0.27	0.01	0.24	0.01	-0.15	0.18	0.89	0.18
Schoolwork impulsivity	-0.30	0.00	0.12	0.00	-0.17	0.10	0.75	0.10
Interpersonal impulsivity	-0.19	0.05	0.59	0.05	-0.09	0.35	0.96	0.36
Grit score	0.16	0.09	0.76	0.10	0.16	0.13	0.82	0.13
Growth mindset	0.15	0.19	0.91	0.18	0.22	0.03	0.46	0.04
Affective empathy (Contagion)	-0.04	0.55	0.99	0.55	-0.04	0.62	0.99	0.62
Cognitive empathy (Understanding)	-0.07	0.35	0.96	0.36	-0.05	0.41	0.97	0.40
Prosocial motivation (Support)	0.05	0.50	0.99	0.49	0.05	0.51	0.99	0.51
Leadership	0.12	0.20	0.90	0.20	0.14	0.07	0.65	0.07
Communication	0.10	0.35	0.96	0.35	0.03	0.80	0.99	0.80
Planning	0.05	0.65	0.99	0.64	0.03	0.75	0.99	0.74
Emotional symptoms	-0.17	0.02	0.31	0.02	-0.13	0.07	0.65	0.07
Conduct problem	-0.21	0.01	0.16	0.01	-0.13	0.10	0.77	0.10
Hyperactivity	-0.20	0.00	0.13	0.00	-0.02	0.82	0.99	0.82
Peer problem	-0.17	0.01	0.24	0.01	-0.09	0.21	0.92	0.21
Prosocial	0.24	0.00	0.08	0.00	0.12	0.11	0.74	0.11

Note: This table exhibits the effect of the treatment on the standardized outcome variables. All outcome variables are standardized [(y_i – mean of the control group)/standard deviation of control group]. Coefficients are estimated with OLS. Baseline controls included: children's age, gender, baseline literacy score, baseline numeracy score, access to private tuition, parents' education in years, family income, religion, access to TV & smartphone, homestead size and the number of members in the household. Children's grade fixed effects are included in all regressions. Westfall-Young FWER adjusted p-values and Randomized Inference (RI) P-values are calculated based on 5,000 replications.

Table 5. Effects of dosage on learning outcomes.

Variables	Standardized coefficient on dosage	
	(1) OLS	(2) IV
Total score	0.01***	0.02***
	(0.00)	(0.00)
Literacy	0.01***	0.03***
	(0.00)	(0.00)
Numeracy	0.01***	0.03***
	(0.00)	(0.00)

Note: This table exhibits the effect of the dose of the intervention on the standardized learning outcome variables. Only literacy and numeracy lessons are considered in calculating dose. In the first row, all literacy and numeracy lessons are used in calculating dose as the outcome is total score. In the second row, only literacy lessons are used in calculating dose as the outcome is literacy score. In the bottom row, only numeracy lessons are used in calculating dose as the outcome is numeracy score. In Column 1, estimates from OLS regressions are exhibited where the key explanatory variable is the number of relevant lessons taken. In Column 2, estimates from instrumental variable regressions are exhibited where treatment assignment serves as the instrument for the number of relevant lessons taken. Baseline controls included: children's age, baseline literacy score, baseline numeracy score, access to private tuition, parents' education in years, family income, religion, access to TV & smartphone, homestead size and the number of members in the household. Children's grade fixed effects are included in all regressions. Robust standard errors clustered at the village level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6. Mean and difference of homeschooling outcomes, by treatment.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	T1: Standard	T2: Extended	T3: Control	Difference (T1 & T3)	Difference (T2 & T3)	Difference (T1 & T2)	Hypothesized sign
Student's study time	92.06	90.36	87.03	5.03	3.32	1.71	(+)
	(3.53)	(3.26)	(3.70)	(5.13)	(4.94)	(4.81)	
				[5.78%]	[3.82%]	[1.89%]	
Extent of study	2.83	2.86	2.65	0.19**	0.21***	-0.03	(+)
	(0.06)	(0.06)	(0.04)	(0.07)	(0.08)	(0.09)	
				[7.00%]	[8.02%]	[-0.94%]	
Caregiver's time in homeschooling	85.92	78.60	75.48	10.44**	3.13	7.32**	(+)
	(2.68)	(2.04)	(3.41)	(4.34)	(3.98)	(3.37)	
				[13.83%]	[4.14%]	[9.31%]	
Extent of caregiver's homeschooling	2.87	2.76	2.63	0.24***	0.13	0.11	(+)
	(0.05)	(0.05)	(0.07)	(0.09)	(0.08)	(0.07)	
				[9.00%]	[4.76%]	[4.05%]	
Observation	567	562	561	1128	1123	1129	-
Village	30	30	30	60	60	60	-

Note: Student's study time – daily study time in minutes provided by the children; Extent of study – a 5-point Likert-scale response; 'none' to 'a great deal'; Caregiver's time in homeschooling – daily homeschooling time provided by the caregivers in minutes; Extent of caregiver's homeschooling – a 5-point Likert-scale response; 'none' to 'a great deal'. In columns 1-3, the means of each of the three treatment arms are presented. Columns 4-6 present estimates of univariate OLS regressions, where the omitted treatment category is T3 in column 4, T3 in column 5, and T2 in column 6. Column 7 exhibits the hypothesized direction of the differences in columns 4 and 5. There was no a priori assumption about the difference between the two treatment arms. Robust standard errors clustered at the village level are in parentheses. Percentages of change between groups marked in column-header are in square brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Treatment effect on homeschooling outcomes.

Variables	T1: Standard				T2: Extended			
	(1) Coefficient on treatment dummy	(2) t-test *P-value*	(3) FWER *P-value*	(4) RI *P-value*	(5) Coefficient on treatment dummy	(6) t-test *P-value*	(7) FWER *P-value*	(8) RI *P-value*
Student's study time	0.10	0.26	0.68	0.26	0.07	0.43	0.68	0.42
Extent of study	0.20	0.01	0.11	0.01	0.24	0.00	0.04	0.00
Caregiver's time in home-schooling	0.20	0.03	0.21	0.03	0.07	0.39	0.68	0.40
Extent of caregiver's homeschooling	0.23	0.01	0.08	0.01	0.13	0.12	0.44	0.11

Note: This table exhibits the effect of the treatment on the standardized outcome variables. All outcome variables are standardized [(y_i – mean of the control group)/standard deviation of control group]. Coefficients are estimated with OLS. Baseline controls included: children's age, gender, baseline literacy score, baseline numeracy score, access to private tuition, parents' education in years, family income, religion, access to TV & smartphone, homestead size and the number of members in the household. Children's grade fixed effects are included in all regressions. Westfall-Young FWER adjusted p-values and Randomized Inference (RI) P-values are calculated based on 5,000 replications.

Table 8. Heterogenous treatment effect on learning and homeschooling

Dependent variables	**W: Gender**			**X: Baseline score**			**Y: Household income**			**Z: Parental education**		
	(1) Boy	(2) Girl	(3) Inter-action	(4) Above median	(5) Below median	(6) Inter-action	(7) Above median	(8) Below median	(9) Inter-action	(10) Above median	(11) Below median	(12) Inter-action
Panel A: Standard treatment												
Total score	0.55***	0.65***	-0.09	0.48***	0.75***	-0.25**	0.57***	0.65***	-0.08	0.53***	0.75***	-0.21**
	(0.09)	(0.10)	(0.10)	(0.10)	(0.09)	(0.11)	(0.08)	(0.11)	(0.11)	(0.08)	(0.11)	(0.10)
Literacy	0.54***	0.57***	-0.02	0.48***	0.65***	-0.17	0.50***	0.64***	-0.13	0.51***	0.64***	-0.13
	(0.10)	(0.10)	(0.10)	(0.11)	(0.09)	(0.12)	(0.09)	(0.11)	(0.11)	(0.08)	(0.12)	(0.11)
Numeracy	0.43***	0.62***	-0.18*	0.38***	0.73***	-0.32**	0.54***	0.51***	0.02	0.42***	0.72***	-0.29***
	(0.09)	(0.08)	(0.10)	(0.08)	(0.09)	(0.10)	(0.08)	(0.09)	(0.11)	(0.07)	(0.09)	(0.09)
Student's study time	0.10	0.11	-0.02	0.13	0.04	0.08	0.04	0.18	-0.14	0.11	0.13	0.00
	(0.11)	(0.12)	(0.14)	(0.11)	(0.10)	(0.12)	(0.10)	(0.12)	(0.13)	(0.11)	(0.14)	(0.15)
Caregiver's time in homeschooling	0.22**	0.19*	0.04	0.24**	0.16	0.08	0.21**	0.17	0.02	0.35***	0.03	0.30**
	(0.11)	(0.11)	(0.13)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)	(0.12)	(0.10)	(0.14)	(0.14)
Panel B: Extended treatment												
Total score	0.61***	0.69***	-0.09	0.60***	0.69***	-0.10	0.52***	0.78***	-0.26**	0.56***	0.77***	-0.22**
	(0.08)	(0.09)	(0.10)	(0.09)	(0.10)	(0.10)	(0.07)	(0.10)	(0.10)	(0.08)	(0.10)	(0.10)
Literacy	0.60***	0.63***	-0.03	0.61***	0.61***	-0.00	0.46***	0.76***	-0.29***	0.53***	0.72***	-0.20*
	(0.09)	(0.09)	(0.10)	(0.09)	(0.10)	(0.11)	(0.08)	(0.11)	(0.10)	(0.09)	(0.10)	(0.11)
Numeracy	0.48***	0.62***	-0.15	0.44***	0.65***	-0.22**	0.49***	0.62***	-0.14	0.47***	0.66***	-0.19*
	(0.08)	(0.08)	(0.10)	(0.08)	(0.09)	(0.09)	(0.07)	(0.09)	(0.10)	(0.07)	(0.09)	(0.10)
Student's study time	0.14	0.02	0.11	0.04	0.12	-0.07	0.01	0.16	-0.16	0.12	0.03	0.12
	(0.12)	(0.10)	(0.13)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)	(0.12)	(0.10)	(0.11)	(0.12)
Caregiver's time in homeschooling	0.03	0.12	-0.05	0.15	-0.00	0.13	0.07	0.08	-0.03	0.07	0.08	0.00
	(0.09)	(0.10)	(0.13)	(0.09)	(0.10)	(0.10)	(0.09)	(0.11)	(0.11)	(0.09)	(0.11)	(0.11)

Note: This table exhibits the heterogeneous treatment effects of the intervention on the learning outcome and homeschooling variables. Coefficients are estimated with OLS. The dependent variable for each regression is listed in the first row. The same list of control variables is used as before. Children's grade fixed effects are included in all regressions. Boy = dummy variable for boy participant; above-median = dummy (1 if the corresponding value is above the median); interaction = interaction term between treatment and gender or above median variable. Robust standard errors clustered at the village level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendices

Appendix A: Additional Material

A-1.Study context and sample A3
A-1.1. Baseline sample A4
A-2.Interactive Audio Instructions (IAI) method A4
A-2.1. Literacy module A5
A-2.2. Numeracy module A5
A-2.3. Leadership module A6
A-3.Learning content of different modules A7
A-3.1. Literacy module A7
A-3.2. Numeracy module A9
A-3.3. Leadership module A12
A-4.Outcomes and instruments for measurement A13
A-4.1. Learning outcomes A13
A-4.2. Leadership, communication, and planning skills A14
A-4.3. Noncognitive skills A15
A-4.4. Behavioral difficulties A17
A-4.5. Homeschooling time (student) A18
A-4.6. Homeschooling time (Mother or caregiver) A18
A-5.Social desirability bias A19
A-6.Assessment test reliability A19

Appendix B: Additional Figures and Tables

Figures

Figure B1. IVR flow diagram of T1: Standard treatment group. A21
Figure B2. IVR flow diagram of T1: Extended treatment group. A22
Figure B3. Number of lessons listened by caregiver-child dyads, by module. A23
Figure B4. Attendance of caregiver-child dyads, by lesson. A24
Figure B5. Assessment-test score distribution, by treatment. A25
Figure B6. Differential Item Functioning (DIF). A26

Tables

Table B1. Survey attrition rate. A27
Table B2. Attrition, by treatment. A28
Table B3. Percentage of children who answered correctly, by question and treatment. A29
Table B4. Parents' social desirability bias analysis. A30
Table B5. Child's social desirability bias analysis. A31
Table B6. Reliability of endline assessment test questions. A32
Table B7. Differential Item Functioning (DIF) analysis. A33
Table B8. Differential Item Functioning (DIF) adjusted learning outcome. A34

Appendix A: Additional Material

A-1. Study context and sample

In partnership with the Global Development and Research Initiative (GDRI) — a local non-governmental organization (NGO) in Bangladesh, we worked with 1,763 primary school-aged children and their caregivers in 90 villages in two southwestern districts (Khulna and Satkhira). GDRI had worked with more than 7,500 children across 223 villages in these areas before the Covid-19 pandemic.[1] We, therefore, had access to household contact information and pre-pandemic learning levels of the children.

As of January 1, 2021, the average age of the children we worked with was 7.4 years and the age range was 5.1 to 9.9 years. These children were from households with a low socio-economic status. The average monthly income per household was BDT 11,008 (USD 130.8) in 2021 whereas the national average monthly income per rural household was BDT 13,998 (USD 166.3) back in 2016 (BBS, 2017).[2] In terms of access to distance learning modalities, these households were also quite disadvantaged. In May 2021, only 46.5% of them had access to TV, and less than 1% had access to computers or radio. All of them had access to mobile phones, 35.6% of which were smartphones. However, these smartphones were rarely used for educational purposes due to the high data costs and low internet speed, particularly in rural areas. The caregivers of these children also had low levels of education; the average years of schooling for the fathers and mothers was 5.9 and 7.1 years, respectively. Overall, dues to unavailability of Information and Communication Technologies (ICT) and lower caregiver capacity, these children did not have access to most distance learning modalities and received poor homeschooling support.

In this context, feature phones could provide access to distance learning and could assist caregivers to engage in their children's education. However, content that could be broadcast via feature phones was limited to audio or short phone mobile messages. Therefore, to improve the efficiency of audio content, we incorporated Interactive Audio Instruction (IAI) to turn a one-way technology into a tool for active learning (Bosch et al., 2002). Though this method was originally developed for radio platforms, we adopted it for Interactive Voice Response (IVR).

[1] These children were part of a completed project named Investing in our future by GDRI.
[2] Bangladeshi Taka — currency of Bangladesh. As of December 1, 2021, 1 USD = 84.1872x.

A-1.1. Baseline sample

From the list of contacts provided by GDRI, we, first, randomly selected 90 villages and then 3,000 households with mobile phone numbers. We were able to reach and complete a baseline survey for 2,400 children from 2,387 households. Others did not respond, or the phone was not active, invalid, or switched off, or they were not interested in this intervention. We randomly selected about 16–22 children[3] from each of these 90 villages and randomly assigned the villages to one of three treatment arms, i.e., Standard, Extended, and Control. Our final sample size was 1,763 children from 1,755 households at the baseline. After randomization, we checked whether several socio-economic characteristics and children's assessment scores were balanced across the treatment arms. These results are shown in Table. These characteristics are balanced across the treatment and control groups.

A-2. Interactive Audio Instructions (IAI) method

The original design of Interactive Audio Instruction was created back in 1970 as Interactive Radio Instructions (IRI) to teach mathematics via radio in Nicaragua (Bosch, 1997). IAI is an instructional approach that turns a one-way technology into a tool for active learning as it requires learners to stop and react to questions and exercises through verbal response, to engage in group work, and physical and intellectual activities while the program is on the air (Bosch et al., 2002). Facilitators play an important role in IAI in supervising the progression of the lessons (Ho & Thukral, 2009). IAI generally follows the constructivist approach of teaching whereby children develop their imagination and make their own knowledge, which is determined by their own experiences (Elliott, Kratochwill, & Travers, 2000).

Just before the intervention, GDRI field staff visited the sample households to provide a guidebook and briefing on the IAI method. They particularly explained how interaction will happen, what will be the role of caregivers, and what are the things caregivers will need during the lesson play. Caregivers or mothers were the facilitators in our intervention. For example, they may draw a few figures, show some letters, or write numbers to engage with their learners during the IAI sessions played over the IVR. In the following subsections we provide a few such examples or activities.

[3] There are four villages with less than 16 children: 10 children each in two villages, 11 children in one village and 13 children in one village. We capped the sample size at 22 children per village to match the budgetary allocation of this study.

A-2.1. Literacy module

For literacy lessons, caregivers or mothers were given alphabet sheets (illustrated below) to practice or to follow the instructions given by the instructor during the audio lessons played over IVR. Also, we relied on the Bangla textbook provided by the government for them to practice various literacy exercises.

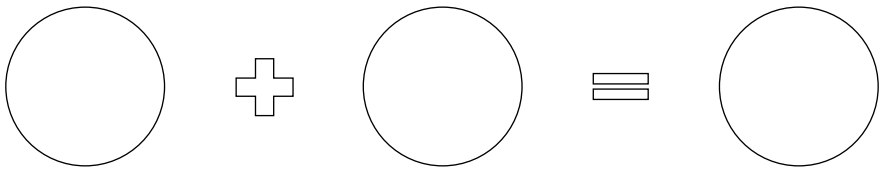
English Alphabet

Bangla Alphabet

A-2.2. Numeracy module

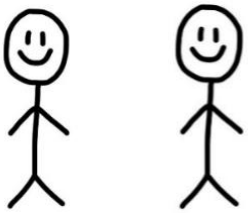
Counter: A counter is anything we can count. For example, sticks, marbles, pieces of bricks or colored pencils. The caregiver had to arrange the counters before the lesson and use them to help the child practice counting during the numeracy lessons.

Number partner: A number partner is two numbers when added together we get another new number. For example, 5 plus 5 equals 10. In this case, 5 and 5 are the number partners of 10. Similarly, 7 and 3 also add up to 10, so these two numbers are also number partners. A number can have many number partners. For example, the caregiver had to draw figures like the following one in a few lessons to help the child identify the number partners.



A-2.3. Leadership module

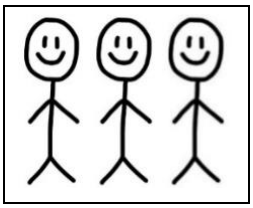
Stick figure: Stick figures are graphical representations of humans used for explaining various contexts to children. In leadership lessons, there were various situations where the child had to decide as per instructions given by the instructor during the audio lessons. For example, the caregiver had to draw stick figures to visually explain the situations, as follows:



Oral saline: In lesson 5 of the leadership module, the caregiver taught their child how to make oral saline. The recipe for homemade oral saline was given to the caregiver beforehand in the guidebook. For example, *"The process of making oral saline is very simple and you all know it. 1 handful of sugar or molasses, 3 finger pinches of salt mixed with half a liter of clean water to make oral saline."*

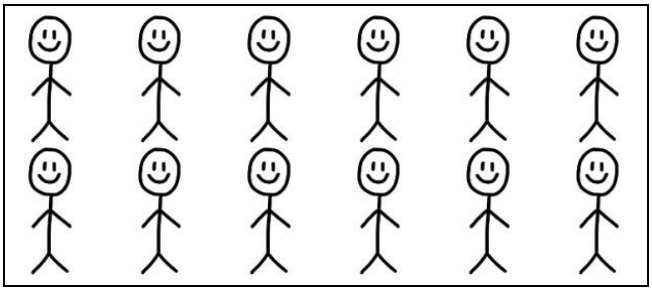
Family picture: In a few lessons, the caregiver had to draw a family figure. During the lesson, the instructor gave instructions to the caregiver to interact with their child using the family figure. A sample family figure is given as follows:

Family



Classroom: The caregiver had to draw a classroom picture on a few occasions. An example of this picture is as follows:

Classroom and classmates



A-3. Learning content of different modules

The literacy and numeracy modules were developed based on the lessons of the 'Rising on Air' program and the leadership module was developed based on the 'LEAD Learning' program. These lessons were completely rewritten and contextualized for Bangladesh by local educators and educational researchers. These contents are briefly presented in the following sections.

A-3.1. Literacy module

Brief contents of the literacy module are exhibited in the following table:

Table A1. Literacy module's lesson plan

No	Competency	Content	Sub-content	Intervention lesson	Learning outcome
1	Listening; Speaking; Reading.	Sentence; Words; Story	Sentence Warm-up	Rice is healthy	Student will be able to (SWBAT) understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Boat; Cat	
			Brilliant Blending Words	Pen; Man; Rat; Nap	
			Story	The Lost Laugh (1)	
2	Listening; Speaking; Reading.	Sentence; Words; Story	Sentence Warm-up	Cats have powerful eyes	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Tongue; Kitchen; Bird	
			Brilliant Blending Words	Pin; Fun; Tin	
			Story	The Lost Laugh (2)	
3	Listening; Speaking; Reading.	Sentence; Words; Story	Sentence Warm-up	A mosquito is an insect	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Wave; Beach; Sky	
			Brilliant Blending Words	Pet; But; Wet; Map	
			Story	The Bee & the Elephant	
4	শোনা, বলা, পড়া, লিখা *(In Bangla)*	নতুন বাংলা শব্দ; সমার্থক শব্দ; বিপরীতার্থক শব্দ	ছড়া – শব্দ গঠন	আতা গাছে তোতা পাখি	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	পাখি- পক্ষী, বিহগ, খেচর গাছ- বৃক্ষ, তরু, শাখী	
			বিপরীত শব্দ	ছোট- বড়, জয়-পরাজয়, বর-বধূ, বেশি -কম	
			গল্পের রাজ্য	কাক ও কলসি	
			নতুন শব্দ ও বাক্য গঠন	নুড়ি, কাক, কলসি	
5	Listening; Speaking; Reading.	Sentence; Words; Story	Sentence Warm-up	Babies are born without teeth	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Three; Door; Rice	
			Brilliant Blending Words	Fish; Path; Bath	
			Story	The Toothache (1)	
6	Listening; Speaking; Reading.	Sentence; Words; Story	Sentence Warm-up	My dad is caring	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Hat; Fan; Bag	
			Listen and Write Words	Cat; Van; Man	
			Story	The Toothache (2)	
7	Listening; Speaking; Reading.	Sentence; Words; Story	Sentence Warm-up	Clouds are made up of tiny droplets of water	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Win; Sing; Dress	
			Listen and Write Words	Chop; Sad; Rock	
			Story	The Red Raincoat (1)	
8	শোনা, বলা, পড়া, লিখা *(In Bangla)*	বাংলা নতুন বাংলা শব্দ, সমার্থক ও বিপরীতার্থক শব্দ	ছড়া	আমি হব-কাজী নজরুল ইসলাম	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	রাত- রত্রি, নিশি, রজনী সূর্য- রবি, তপন, ভানু	
			বিপরীত শব্দ	আলসে-কর্মঠ, সকাল-বিকাল, আগে-পরে, ঘুমিয়ে-জেগে	
			গল্পের রাজ্য	দুখুর ছোটবেলা	
			নতুন শব্দ ও বাক্য গঠন	ঝাঁকড়া, মস্তব, জাতীয়	
9	Listening; Speaking; Reading; Writing	Sentence; Words; Story	Sentence Warm-up	Fruit and vegetables are good for you	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Pin; Pen; Pan	
			Brilliant Blending Words	Beg; Fish; Ten	
			Story	The Red Raincoat (2)	
10	Listening; Speaking; Reading; Writing	Sentence; Words; Story	Sentence Warm-up	Buses carry many passengers	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Bit; Red; Ten	
			Brilliant Blending Words	Fill; Hill; Sand	
			Story	The Boy's New Bike (1)	

11	Listening; Speaking; Reading; Writing		Sentence Warm-up	When you are honest you are telling the truth	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Brilliant Blending Words	Crop; Swim; Skip	
			Listen and Write Words	Crop; Swim; Skip	
			Story	The Boy's New Bike (1)	
12	শোনা, বলা, পড়া, লিখা *(In Bangla)*	বাংলা নতুন বাংলা শব্দ, সমার্থক ও বিপরীতার্থক শব্দ	ছড়া -শব্দ গঠন	কাজের আনন্দ- নবকৃষ্ণ ভট্টাচার্য আহরণ, কব, তৃণলতা, বুনি, সঞ্চয়, পিপীলিকা	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	ভাই- ভ্রাতা, ভাইয়া, সহোদর পিপীলিকা- পিঁপড়া, পিপড়ে	
			বিপরীত শব্দ	শীত- গ্রীষ্ম, আসা- যাওয়া, দাঁড়ানো - বসা, আনা- নেওয়া	
			গল্পের রাজ্য	শীতের সকাল	
			নতুন শব্দ ও বাক্য গঠন	পোহান, নাশতা	
13	Listening; Speaking; Reading. Writing		Sentence Warm-up	The girl was crying because she was sad	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Brilliant Blending Words	Ask; Text; Next	
			Listen and Write Words	Ask; Text; Desk	
			Story	My Feelings (1)	
14	Listening; Speaking; Reading. Writing		Sentence Warm-up	Just like humans, plants need water to survive	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Brilliant Blending Words	Grand; Crept; Past	
			Listen and Write Sentence	A thin rat crept past the fat cat	
			Story	My Feelings (2)	
15	Listening; Speaking; Reading. Writing		Sentence Warm-up	Plants can be used as medicine	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Brilliant Blending Words	Fish; Stretch; Net	
			Listen and Write Sentence	The men on the ship stretch a big fishing net	
			Story	The Red Plant (1)	
16	শোনা, বলা, পড়া, লিখা *(In Bangla)*	বাংলা নতুন বাংলা শব্দ, সমার্থক ও বিপরীতার্থক শব্দ	ছড়া -শব্দ গঠন	ট্রেন- শামসুর রাহমান	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	নদী- তটিনী, তরঙ্গিনী, প্রবাহিণী বাড়ি- গৃহ, আবাস, নিবাস	
			বিপরীত শব্দ	ঘরে- বাইরে, দেশ- বিদেশ, ছোটা- থামা	
			গল্পের রাজ্য	গাছ লাগানো	
			নতুন শব্দ ও বাক্য গঠন	গাছ	
17	Listening; Speaking; Reading.		Sentence Warm-up	It is important to exercise every day	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Cat/Cart; Pot/Port; Shirt/Shit	
			Brilliant Blending Words	Sit; Birth; Torn	
			Story	The Red Plant (2)	
18	Listening; Speaking; Reading.		Sentence Warm-up	Keep your mouth happy by brushing your teeth	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Gem/Germ; Hut/Hurt	
			Brilliant Blending Words	Ten; Term; Turn; Sun	
			Story	The Magical Lunchbox (1)	
19	Listening; Speaking; Reading.		Sentence Warm-up	Colors help you remember memories	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Sport; Spurt; Barn	
			Listen and Write Words	Corn; planner	
			Story	The Magical Lunchbox (2)	
20	শোনা, বলা, পড়া, লিখা *(In Bangla)*	বাংলা নতুন বাংলা শব্দ, সমার্থক ও বিপরীতার্থক শব্দ	ছড়া -শব্দ গঠন	বড় কে- হরিশ্চন্দ্র মিত্র	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	সংসার- জগত, ভুবন, বিশ্ব বড়- বৃহৎ, প্রকাণ্ড	
			বিপরীত শব্দ	বড়- ছোট, সহজ - কঠিন, যার - তার, সুখ - দুঃখ, দোষ- গুণ	
			গল্পের রাজ্য	পাখিদের কথা	
			নতুন শব্দ ও বাক্য গঠন	সংসার; কঠিন; ব্যাপার প্রিয় পাখি	
21	Listening; Speaking; Reading. Writing		Sentence Warm-up	Planting trees is good for the environment	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Write and Read Words	Strong; Swimmer; Faster; Longer	
			Listen and Write Sentence	That girl is a strong swimmer. She can swim faster and longer than me	
			Story	Water is Important (1)	
22	Listening; Speaking; Reading.		Sentence Warm-up	Drinking water is needed for humans to stay alive	SWBAT understand English sentence structure; basic
			Sound Workout Words	Bit; Bite	

	Writing		Brilliant Blending Words	Rip; Ripe; Hid; Hide	spelling; and the meaning of a short story.
			Story	Water is Important (1)	
23	Listening; Speaking; Reading. Writing		Sentence Warm-up	People from different countries can wear different clothes	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	June; But	
			Listen and Write Words	Bone; Kite; Sharp	
			Story	The Hat Seller (1)	
24	শোনা, বলা, পড়া, লিখা *(In Bangla)*	বাংলা নতুন বাংলা শব্দ, সমার্থক ও বিপরীতার্থক শব্দ	ছড়া -শব্দ গঠন	মামার বাড়ি- জসীমউদদীন	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	ঝড়- ঝঞ্ঝা, ঝটিকা, ঝাপ্টা, পুষ্প- কুসুম, ফুল, কুঁড়ি	
			বিপরীত শব্দ	কাঁচা- পাকা, উঠা- নামা, ছেলে- বুড়ো,	
			গল্পের রাজ্য	পিঁপড়ে ও ঘুঘু	
			বাক্য গঠন	পিঁপড়ে ও ঘুঘু	
25	Listening; Speaking; Reading. Writing		Sentence Warm-up	Fingernails grow faster than toenails.	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Write and Read Words	Read; Tale; Page	
			Listen and Write Sentence	We like this book. It has tales.	
			Story	The Hat Seller (2)	
26	Listening; Speaking; Reading. Writing		Sentence Warm-up	People who watch sports can be called spectators	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Plate; Play; Rain	
			Write and Read Words	Train; Tray	
			Story	The Lost Ball	
27	শোনা, বলা, পড়া, লিখা *(In Bangla)*	বাংলা নতুন বাংলা শব্দ, সমার্থক ও বিপরীতার্থক শব্দ	ছড়া -শব্দ গঠন	আদর্শ ছেলে	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	মানুষ- মানব, নর, লোক ভয়- ডর, ভীতি, আতঙ্ক	
			বিপরীত শব্দ	হাত- পা, বিপদ- আপদ, হাসি- কান্না, কল্যাণ- অকল্যাণ	
			গল্পের রাজ্য	একজন পটুয়ার কথা	
			বাক্য গঠন	কামরুল হাসান	
28	Listening; Speaking; Reading. Writing		Sentence Warm-up	A triangle is a shape with three sides	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Sound Workout Words	Kite; Child; Sky	
			Write and Read Words	Fit; Fried; Night	
			Story	When Will Mother Be Back? (1)	
29	Listening; Speaking; Reading. Writing		Sentence Warm-up	You can use your fingers to count to ten	SWBAT understand English sentence structure; basic spelling; and the meaning of a short story.
			Spelling Sound out	Glue; Fruit; New	
			Write and Read Words	Cup; Value; Juice	
			Story	When Will Mother Be Back? (2)	
30	শোনা, বলা, পড়া, লিখা *(In Bangla)*	বাংলা নতুন বাংলা শব্দ, সমার্থক ও বিপরীতার্থক শব্দ	ছড়া -শব্দ গঠন	আমাদের এই বাংলাদেশ- সৈয়দ শামসুল হক	SWBAT understand the synonym and antonym of Bengali words and the meaning of a short story.
			প্রতিশব্দ	দেশ- রাষ্ট্র, স্বদেশ, জন্মভূমি বীর- সাহসী, নির্ভীক, অকুতোভয়	
			বিপরীত শব্দ	পূর্ব- পশ্চিম, স্বাধীন- পরাধীন, আপন- পর, বীর- কাপুরুষ, প্রিয়- অপ্রিয়	
			গল্পের রাজ্য	ভাষাশহিদের গল্প	
			বাক্য গঠন	মাতৃভাষা	

A-3.2. Numeracy module

Brief contents of the numeracy module are exhibited in the following table:

Table A2. Numeracy module's lesson plan

No	Competency	Content	Sub-content	Intervention lesson	Learning outcome
1	Counting	Counting 1 to 10; Identification 1 to 10	Mindful moment	Counting with breathing	Students will be able to (SWBAT) count out and write represent a quantity 1-10; SWBAT identify numerals 1-10
			Number warmup	Counting	
			Number workout	Counters	
			Brain break	Right-left	
			Challenge problem	Hand	
2			Mindful moment	Music	

No	Competency	Content	Sub-content	Intervention lesson	Learning outcome
	Counting; Addition	Counting 1 to 20; simple addition	Number warmup	Counting	SWBAT count out and write represent a quantity 1-20; SWBAT identify numerals 1-20; SWBAT add simple numbers
			Number workout	Compare	
			Brain break	Music	
			Challenge problem	Addition	
3	Addition	Simple two numbers addition	Mindful moment	Sound focus	SWBAT solve Joining problems (using counters, counting all, or counting on) and write a matching equation.
			Number warmup	Count and move	
			Number workout	Addition	
			Brain break	Action	
			Challenge problem	Count	
4	Subtraction	Simple two numbers subtraction	Mindful moment	Balance exercise	SWBAT solve separating problems (using counters, counting all, or counting on) and write a matching equation when the part is unknown.
			Number warmup	Counting backwards	
			Number workout	Subtraction	
			Brain break	Dance	
			Challenge problem	Subtraction	
5	Addition and subtraction	Solve addition and subtraction problems	Mindful moment	Muscle squeeze exercise	SWBAT represent and solve addition and subtraction problems.
			Number warmup	Count and move	
			Number workout	Addition and subtraction	
			Brain break	Shakedown	
			Challenge problem	Addition & subtraction	
6	Number Comparison	Decomposition, Addition, Subtraction and Comparison	Mindful moment	Deep breathing exercise	SWBAT review the concept of Writing numbers, Decomposition, Addition, Subtraction and Comparison of numbers from 1 to 10
			Number warmup	Count and move	
			Number workout	Comparison	
			Brain break	Role play	
			Challenge problem	Addition	
7	Ordinal Number	Ordinal number	Mindful moment	Balance exercise	SWBAT identify and describe ordinal numbers (first to tenth)
			Number warmup	Count and move	
			Number workout	Ordinal number	
			Brain break	Song	
			Challenge Problem	tricky numbers	
8	Numerals	different sets of groups, number bonds and write a matching number sentence	Mindful Moment	Sound Focus exercise	SWBAT determine and write how many objects are in a set (10-15) and identifies numerals 10-15
			Number Warmup	Detectives	
			Number Workout	Numerals	
			Brain Break	Freeze Dance	
			Challenge Problem	Number Bond	
9	Ordinal numbers	Identify and describe ordinal numbers (eleventh to twentieth)	Mindful Moment	Balance exercise	SWBAT identify and describe ordinal numbers (eleventh to twentieth)
			Number Warmup	Skip counting	
			Number Workout	Ordinal Numbers	
			Brain Break	Freeze Dance	
			Challenge Problem	Identifying circle	
10	Teen numbers	Understand teen numbers are composed of tens and ones	Mindful Moment	Sound focus exercise	SWBAT understand, identify, and write teen numbers
			Number Warmup	Guess number	
			Number Workout	Teen numbers	
			Brain Break	As If	
			Challenge Problem	Teen numbers	
11	Multiple addition equations	Understand and apply multiple addition equations	Mindful Moment	Balance exercise	SWBAT add more than two numbers and write the equation in multiple ways
			Number Warmup	Beep counting	
			Number Workout	Commutative property of addition	
			Brain Break	Creative dance	
			Challenge Problem	True and false	
12	Subtraction	subtracting two values using different ways	Mindful Moment	Sound sense	SWBAT subtract two values using fingers, counting back, counting up, or known facts.
			Number Warmup	count backwards	
			Number Workout	different ways of subtracting two values	
			Brain Break	Dance party	
			Challenge Problem	Subtraction problems	
13	addition and subtraction	Solving addition and subtraction problems	Mindful Moment	muscle squeeze exercise	SWBAT solve addition and subtraction problems using ten as a landmark.
			Number Warmup	Count and move	
			Number Workout	Solving addition and subtraction problems together	
			Brain Break	Shakedown	
			Challenge Problem	Mystery Number	
14	addition and subtraction comparison	Solving addition and subtraction comparison problems	Mindful Moment	Balancing exercise	SWBAT solve comparison problems using addition and subtraction.
			Number Warmup	Count and move	
			Number Workout	Comparison addition & Subtraction	
			Brain Break	Creative Dance	

No	Competency	Content	Sub-content	Intervention lesson	Learning outcome
			Challenge Problem	Number sentence	
15	Equal sign	Understanding equal sign and comparing both sides of the equal sign	Mindful Moment	Balance exercise	SWBAT understand the meaning and application of equal sign
			Number Warmup	Beep Counting	
			Number Workout	Identifying True-False	
			Brain Break	Alphabet Mov	
			Challenge Problem	Distinguishing between true-false equations	
16	Expanded form of Numbers	Understand and write the expanded or standard form of numbers	Mindful Moment	Sounds workout	SWBAT represent numbers 10-99 in expanded form.
			Number Warmup	Count forward by tens	
			Number Workout	Standard form	
			Brain Break	As If…	
			Challenge Problem	Mystery Number	
17	Place value to compare two-digit	learn to use place value to help us compare two-digit	Mindful Moment	deep breathing exercise	SWBAT compare two-digit numbers by reasoning about tens and ones.
			Number Warmup	count forward by 2s	
			Number Workout	Comparing numbers	
			Brain Break	As If….	
			Challenge Problem	counting and addition, Mystery number- L.P 11.5	
18	Number & Place value	Digit, Number, Place value, comparison	Mindful Moment	balancing exercise	SWBAT model numbers with more than 9 ones
			Number Warmup	Count forward by 2s	
			Number Workout	Make numbers using place value	
			Brain Break	As If…"	
			Challenge Problem	Compare big and small numbers	
19	Different forms of Numbers	Expanded form, Word form, Standard form, Compare numbers	Mindful Moment	muscle squeeze exercise	SWBAT order numbers in different forms
			Number Warmup	count forward by 2s	
			Number Workout	Various forms of numbers	
			Brain Break	As If…	
			Challenge Problem	Comparing Numbers	
20	Composed numbers	Represent numbers with pictures of sticks and dots, Composed numbers	Mindful Moment	Balance Exercise	SWBAT model numbers with tens sticks and counters, pictures of sticks and dots, and/or counting on
			Number Warmup	Skip count forward and backward	
			Number Workout	Composed numbers	
			Brain Break	Freeze Dance	
			Challenge Problem	Expand the form of numbers	
21	Expanded notation (No regrouping)	Addition, number break	Mindful Moment	Deep breathing exercise	SWBAT add 2 two-digit numbers (no regrouping) using pictures and expanded notation
			Number Warmup	Detectives	
			Number Workout	Expanded notation	
			Brain Break	As If…	
			Challenge Problem	Expanded notation	
22	Expanded notation (With regrouping)	Identifying the tens and ones in a two-digit number	Mindful Moment	Sound focus exercise	SWBAT add 2 two-digit numbers (with regrouping) using pictures and expanded notation.
			Number Warmup	Addition	
			Number Workout	Adding numbers	
			Brain Break	Shakedown	
			Challenge Problem	Addition strategies	
23	Subtract (No regrouping)	Addition, Subtraction, Identify the tens and ones in a single- and double-digit number	Mindful Moment	Deep Breathing	SWBAT subtract one- and two-digit numbers from two-digit numbers (no regrouping) using the standard algorithm.
			Number Warmup	Substrate Sentence	
			Number Workout	Subtract	
			Brain Break	Freeze Dance	
			Challenge Problem	subtract with expanded notation	
24	Subtract (With regrouping)	Addition, Subtraction, and multi-step instruction to solve with a standard algorithm	Mindful Moment	Balance Exercise	SWBAT subtract one- and two-digit numbers from two-digit numbers (with regrouping) using the standard algorithm.
			Number Warmup	Skip Count	
			Number Workout	Subtract (with regrouping)	
			Brain Break	Dance Party	
			Challenge Problem	subtract with expanded notation	
25	Addition and Subtraction word problem	Addition, Subtraction	Mindful Moment	Balance Exercise	SWBAT solve 2-digit addition and subtraction word problems using the standard algorithm.
			Number Warmup	2's	
			Number Workout	Addition and Subtraction word problem	
			Brain Break	As if….	
			Challenge Problem	Word Problem	
26	numbers with more than 9 ones	expanded notation, unit form, numbers with stick and dots	Mindful Moment	Sound focus exercise	SWBAT model numbers with more than 9 ones
			Number Warmup	Standard Form	
			Number Workout	Represent Number	
			Brain Break	Dance Party	
			Challenge Problem	Number with unit form	

No	Competency	Content	Sub-content	Intervention lesson	Learning outcome
27	Place value	Greatest to latest, latest to greatest	Mindful Moment	Deep Breath	SWBAT orders 2-digit numbers in different forms
			Number Warmup	Mystery Number	
			Number Workout	Place Value	
			Brain Break	As if….	
			Challenge Problem	Mystery Number	
28	Order Number in a different form	Expanded, standard and written form, largest to smallest and smallest to largest	Mindful Moment	muscle squeeze exercise	SWBAT compare numbers with more than 9 ones
			Number Warmup	Count and move	
			Number Workout	Order number	
			Brain Break	As if….	
			Challenge Problem	Mystery Number	
29	Mental math and Comparison	Greatest to least, least to greatest, smallest to largest, largest to smallest of numbers	Mindful Moment	Balance Exercise	SWBAT review expanded form, +-10/100 mental math, and comparison
			Number Warmup	Count and move	
			Number Workout	Mental math and Comparison	
			Brain Break	Shakedown	
			Challenge Problem	Mystery number	
30	order numbers in many forms	10 ones = 1 ten, 10 tens = 1 hundred, largest to smallest, smallest to largest, solve problems in multiple ways	Mindful Moment	Sound Focus Exercise	SWBAT order numbers in many forms
			Number Warmup	Count and move	
			Number Workout	Number Order	
			Brain Break	Freeze Dance	
			Challenge Problem	Mystery Number	

A-3.3. Leadership module

Brief contents of the leadership module are exhibited in the following table:

Table A3. Leadership module's lesson plan

No	Competency	Content	Sub-content	Intervention lesson	Learning outcome
1	Leadership	What Is Leadership?	Exploration	Leadership; Leader	Introduce students to the concept of leadership
			Blast from the Past	President; Prime Minister; Cricket Captain	
			Creative Corner	Doctor; Nurse	
			Acts of Leadership	2 Hypothetical Stories	
2	Confident; Creative	The 5 Qualities of a Leader, Part 1	Memory Kick	Leadership	Introduce students to the first two qualities – be confident and be creative.
			Exploration	Qualities; Be confident; Be creative	
			Blast from the Past	Gold medal of BD women's cricket team in SAG 2019 (confident) Dr Rafiqul Islam, inventor, ORS (creative)	
			Acts of Leadership	1 Hypothetical Story	
3	Visionary; Teamwork; Delegation of Authority	The 5 Qualities of a Leader, Part 2	Memory Kick	Be confident; Be creative	Introduce students to the last three qualities – set the example, work together, and recognize and applaud success
			Exploration	Qualities; set the example; work together; recognize and applaud success.	
			Blast from the Past	Freedom fighter Mostafa Kamal	
			Creative Corner	Scoring in a football match	
4	Better listening	Active Listening	Memory Kick	set the example; work together; recognize and applaud success	Introduce students to the Importance of listening
			Exploration	Active Listening	
			Blast from the Past	The landscape of the language movement	
			Creative Corner	Father-child conversation	
			Acts of Leadership	1 hypothetical short story	
5	communicate	Communication and Presenting Yourself	Memory Kick	Active Listening	Learn how to communicate well and be approachable
			Exploration	Communication	
			Blast from the Past	Greta Thunberg	
			Creative Corner	Making oral saline	
			Acts of Leadership	1 hypothetical short story	
6	Planning	How to 'Make a Plan'	Memory Kick	Communication	Introduce students to the concept of brainstorming and how to make a plan
			Exploration	Brainstorm; Choose the best idea	
			Blast from the Past	An incident from the independence war	
			Creative Corner	Going to the town	
			Acts of Leadership	1 hypothetical short story	
7	Humble	Bragging vs. Humility	Memory Kick	Planning	
			Exploration	Humility; Bragging	

No	Competency	Content	Sub-content	Intervention lesson	Learning outcome
			Blast from the Past	Muhammad Yunus – Nobel Lecture	Introduce students to the concept of humility and how to avoid bragging
			Acts of Leadership	1 hypothetical short story	
8	Patience	Patience I	Memory Kick	Humility; Bragging	Introduce students to the concept of patience
			Exploration	Patience	
			Blast from the Past	Bangladesh vs Zimbabwe test 2005	
9	Patience	Patience II	Memory Kick	Patience	Example of patience
			Creative Corner	Things to do before play	
			Acts of Leadership	3 hypothetical short stories	
10	Empathy	Empathy I	Memory Kick	Patience	Introduce students to the meaning of empathy and how to be empathetic
			Exploration	Empathy	
			Blast from the Past	Mother Teresa	
11	Empathy	Empathy II	Memory Kick	Empathy	Examples of empathic acts
			Creative Corner	Showing empathy to classmates; Showing empathy to other	
			Acts of Leadership	3 hypothetical short stories	
12	Sympathy; Compassion	Sympathy and Compassion I	Memory Kick	Empathy	Introduce students to the concept of sympathy and how to be compassionate
			Exploration	Sympathy; Compassion	
			Blast from the Past	Rohingya refugees 2017	
13	Sympathy; Compassion	Sympathy and Compassion II	Memory Kick	Sympathy; Compassion	Example of sympathy
			Creative Corner	Prosocial experimental scenario	
			Acts of Leadership	3 hypothetical short stories	
14	Perseverance	Perseverance	Memory Kick	Sympathy; Compassion	Introduce students to the importance of perseverance
			Exploration	Perseverance	
			Blast from the Past	Nelson Mandela [link 1] [link 2]	
			Acts of Leadership	2 hypothetical short stories	
15	Module review	Key Ideas	Review all lessons	Lessons 1 to 14	Review the key ideas of leadership and the qualities of a leader

A-4. Outcomes and instruments for measurement

A-4.1. Learning outcomes

Table A4. Endline assessment test questions

Subject	No	Level 1 (Grade 1)	Level 2 (Grade 2)	Level 3 (Grade 3 & 4)
Literacy	1.	Read aloud the following letters (the first 4 letters from the Bengali alphabet)	Make two words using the Bangla letter ----.	Read aloud this following paragraph (Bangla).
	2.	Fill in the gaps (5 Bangla letters with 2 gaps).	Fill in the gap (a line in Bangla from the textbook)	What is the antonym of the Bangla word (FREEDOM)?
	3.	Make a word with the Bengali letter -----.	What is the spelling of the word (Sundarbans)?	What is the spelling of the word (Bangla of freedom fighter)?
	4.	What is the spelling of (Bengali word)?	What is the antonym of the Bangla word (high)?	What is the meaning of this Bangla word (Bangla word from the textbook)?
	5.	Read the following word (CAP).	Read the following word (FARMER)?	Read aloud this following paragraph (English).
	6.	Answer this English question: What is your name?	Answer this English question: How old are you?	Answer this English question: What month is it now?
	7.	Say the English of Bangla word – (DOOR).	Say the English of Bangla word – (WINDOW).	Say the English of Bangla word – (FARMER).
	8.	Say the English of Bangla word – (BOOK).	Say the English of Bangla word – (UMBRELLA).	Say the English of Bangla word – (WEDNESDAY).
	9.	Say the English of Bangla word – (DOG).	Say the English of Bangla word – (BREAKFAST).	Say the English of Bangla word – (FLAG).
	10.	Spell your name in English.	Read and say the name of these shapes (picture of the square, circle, triangle, and rectangle).	Match the appropriate description with this picture (match from 4 options).

Numeracy	11.	Which number comes after 6? Does it even or odd?	Name the even numbers between 1 and 10.	Sort these three numbers, smallest to the largest (20, 73, 10, 78).
	12.	What is the result of 3+4=?	Sort these three numbers, smallest to largest (23, 17, 38).	There are 6 notes of 20 BDT. How much money is there?
	13.	What is the result of 8-3=?	In a class, there were 16 students. The teacher sends 5 of them for gardening. How many students are left in the classroom?	What is the result of 13+11=?
	14.	How many minutes in 60 seconds?	How many sides a triangle has?	What is the result of 2/4+2/4=?
	15.	What is the result of 6+0=?	There are three fruits on a plate. How many fruits there are in 4 plates?	The price of 5 eggs is BDT 30. How much does it cost to buy 2 eggs?

Note: The test was conducted on a one-on-one basis.

From the answers to these questions, three learning outcome variables were constructed. These are –

- **Total score:** All 15 literary and numeracy questions from Table A4. The test totals eighty points. This variable is continuous; $Y_i^{ts} \in [0,80]$.
- **Literacy:** 10 questions on literacy (English & Bangla). This variable is continuous; $Y_i^{lit} \in [0,50]$.
- **Numeracy:** Five questions on numeracy. This variable is continuous; $Y_i^{num} \in [0,30]$.

A-4.2. Leadership, communication, and planning skills

Table A5. Children's leadership, communication, and planning skills assessment questions

Scale	No	Questions	Answer
[Every statement will start with] My child demonstrates . . .			
7-item Leadership Characteristics	1.	responsible behavior, can be counted on to follow through on activities/projects.	(1) Never (2) Very rarely (3) Rarely (4) Occasionally (5) Frequently (6) Always
	2.	a tendency to be respected by classmates.	
	3.	the ability to articulate ideas and communicate well with others.	
	4.	self-confidence when interacting with age peers.	
	5.	the ability to organize and bring structure to things, people, and situations.	
	6.	cooperative behavior when working with others.	
	7.	a tendency to direct an activity when he or she is involved with others.	
4-item Communication Characteristics	1.	uses voice expressively to convey or enhance meaning.	(1) Never (2) Very rarely (3) Rarely (4) Occasionally (5) Frequently (6) Always
	2.	conveys information nonverbally through gestures, facial expressions, and "body language."	
	3.	is an interesting storyteller.	
	4.	uses colorful and imaginative figures of speech such as puns and analogies.	
15-item Planning Characteristics	1.	determines what information or resources are necessary for accomplishing a task.	(1) Never (2) Very rarely (3) Rarely (4) Occasionally (5) Frequently (6) Always
	2.	grasps the relationship of individual steps to a whole process.	
	3.	allows time to execute all steps involved in a process.	
	4.	foresees consequences or effects of action.	
	5.	organizes his or her work well.	
	6.	takes into account the details necessary to accomplish a goal.	
	7.	is good at games of strategy where it is necessary to anticipate several moves ahead.	
	8.	recognizes the various alternative methods for accomplishing a goal.	

	9.	can pinpoint where areas of difficulty might arise in a procedure or activity.
	10.	arranges steps of a project in a sensible order or time sequence.
	11.	is good at breaking down an activity into step-by-step procedures.
	12.	establishes priorities when organizing activities.
	13.	shows awareness of limitations relating to time, space, materials, and abilities when working on group or individual projects.
	14.	can provide details that contribute to the development of a plan or procedure.
	15.	sees alternative ways to distribute work or assign people to accomplish a task.

From the answers to these questions listed in Table A5, three variables were constructed as follows:

- **Leadership:** Each question of the leadership subscale from the Renzulli scale has a 7-point Likert scoring option. By adding the score of individual questions, the leadership score is calculated. The variable is continuous; $Y_i^{lead} \in [7,42]$.
- **Communication:** Expressive communication subscale from the Renzulli scale has 4 items and each has a 7-point Likert scoring option. The formation of this variable is continuous and calculated as the sum of all sub-questions; $Y_i^{com} \in [4,24]$.
- **Planning:** The planning subscale from the Renzulli scale has 16 questions and each has a similar 7-point Likert scoring option. The variable is continuous; $Y_i^{plan} \in [15,90]$.

A-4.3. Noncognitive skills

Table A6. Children's noncognitive skill assessment survey questions

Scale	No	Questions	Answer
8-item Impulsivity Scale for Children	1.	I forgot something I needed for class.	(1) Almost never (2) About once a month (3) About 2-3 times a month (4) About once a week (5) At least once a day
	2.	I interrupted other students while they were talking.	
	3.	I said something rude.	
	4.	I couldn't find something because my desk, locker, or bedroom was messy.	
	5.	I lost my temper at home or at school.	
	6.	I did not remember what my teacher told me to do.	
	7.	My mind wandered when I should have been listening.	
	8.	I talked back to my teacher or parent when I was upset.	
8-item Grit Scale	1.	New ideas and projects sometimes distract me from previous ones. (R)	(1) Very much like me (2) Mostly like me (3) Somewhat like me (4) Not much like me (5) Not like me at all
	2.	Setbacks don't discourage me.	
	3.	I have been obsessed with a certain idea or project for a short time but later lost interest. (R)	
	4.	I am a hard worker.	
	5.	I often set a goal but later choose to pursue a different one. (R)	
	6.	I have difficulty maintaining my focus on projects that take more than a few months to complete. (R)	
	7.	I finish whatever I begin.	
	8.	I am diligent.	
Growth Mindset Scale	1.	You have a certain amount of intelligence, and you can't really do much to change it.	(1) Strongly agree (2) Agree (3) Mostly agree (4) Mostly disagree (5) Disagree (6) Strongly disagree
	2.	Your intelligence is something about you that you can't change very much.	
	3.	You can learn new things, but you can't really change your basic intelligence.	
EmQue-CA	1.	If my mother is happy, I also feel happy.	(1) Not true (2) Sometimes true (3) often true
	2.	I understand that a friend is ashamed when he/she has done something wrong.	
	3.	If a friend is sad, I like to comfort him.	
	4.	I feel awful when two people quarrel.	

5.	When a friend is angry, I tend to know why.
6.	I would like to help when a friend gets angry.
7.	If a friend is sad, I also feel sad.
8.	I understand that a friend is proud when he/she has done something good.
9.	If a friend has an argument, I try to help.
10.	If a friend is laughing, I also laugh.
11.	If a friend is sad, I understand mostly why.
12.	I want everyone to feel good.
13.	When a friend cries, I cry myself.
14.	If a friend cries, I often understand what has happened.
15.	If a friend is sad, I want to do something to make it better.
16.	If someone in my family is sad, I feel really bad.
17.	I enjoy giving a friend a gift.
18.	When a friend is upset, I feel upset too.

Note: R - reverse scoring.

From the answers to these questions listed in Table A6, multiple outcome variables were constructed:

- **Overall impulsivity:** Each question of the 8-item Impulsivity Scale for Children (ISC) has a 5-point Likert scoring option. Thus, the variable is continuous; $Y_i^{ISC} = \frac{\sum_{q=1}^{q=8} ISC_q}{8} \in [1,5]$.
- **Schoolwork impulsivity:** Items 1, 4, 6, and 7 from the ISC survey are used to construct the schoolwork impulsivity variable. The variable is continuous; $Y_i^{sch} = \frac{\sum_{q=1}^{q=4} ISC_q}{4} \in [1,5]$.
- **Interpersonal impulsivity:** Items 2, 3, 5, and 8 from the ISC survey are used to construct the interpersonal impulsivity variable. The variable is continuous; $Y_i^{per} = \frac{\sum_{q=1}^{q=4} ISC_q}{4} \in [1,5]$.
- **Grit:** Each question of the 8-item grit scale has a 5-point Likert scoring option. The great variable is constructed by averaging these questions. Thus, the variable is continuous; $Y_i^{grit} = \frac{\sum_{q=1}^{q=8} GRIT_q}{8} \in [1,5]$.
- **Growth mindset:** Each question of the 3-item growth mindset scale has a 6-point Likert scoring option. The mean of these questions indicates the growth mindset of the children. Thus, the variable is continuous; $Y_i^{gms} = \frac{\sum_{q=1}^{q=3} GMS_q}{3} \in [1,6]$.
- **Affective empathy (contagion):** 18-item self-report questionnaire that examines the level of empathy in three domains. Each question has a 3-point Likert answering option. All questions are exhibited in Table A6. Items 1, 4, 7, 10, 13, 16, and 18 from the EmQue-CA survey are used to construct the affective empathy of the children. The variable is continuous; $Y_i^{EmA} = \frac{\sum_{q=1}^{q=7} EmQue_q}{7} \in [0,2]$.
- **Cognitive empathy (understanding):** Items 2, 5, 8, 11, and 14 from the EmQue-CA survey are used to construct the cognitive empathy of the children. The variable is continuous; $Y_i^{EmC} = \frac{\sum_{q=1}^{q=5} EmQue_q}{5} \in [0,2]$.

- **Prosocial motivation (support):** Items 3, 6, 9, 12, 15, and 17 from the EmQue-CA survey the EmQue-CA survey are used to construct the cognitive empathy of the children. The variable is continuous; $Y_i^{EmP} = \frac{\sum_{q=1}^{q=6} EmQue_q}{6} \in [0,2]$.

A-4.4. Behavioral difficulties

We used the 25-item parent-report Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). The main motivation for using this scale was to assess whether school closure triggered any fatigue, abnormality or conduct problems among the students and whether participation in the program reduced these problems. Items of this scale are listed in the following Table A7.

Table A7. Children's behavioral difficulties assessment survey questions

No	Questions	Subscale	Answer
Every statement will start with] My child …			
1.	Considerate of other people's feelings	Prosocial	1) Not true 2) Somewhat true 3) Certainly true
2.	Restless, overactive, cannot stay still for long	Hyperactivity	
3.	Often complains of headaches, stomach-aches, or sickness	Emotional Symptoms	
4.	Shares readily with other children, for example toys, treats, pencils	Prosocial	
5.	Often loses temper	Conduct Problem	
6.	Rather solitary, prefers to play alone	Peer Problem	
7.	Generally, well behaved, usually does what adults request (R)	Conduct Problem	
8.	Many worries or often seems worried	Emotional Symptoms	
9.	Helpful if someone is hurt, upset, or feeling ill	Prosocial	
10.	Constantly fidgeting or squirming	Hyperactivity	
11.	Has at least one good friend (R)	Peer Problem	
12.	Often fights with other children or bullies them	Conduct Problem	
13.	Often unhappy, depressed, or tearful	Emotional Symptoms	
14.	Generally liked by other children (R)	Peer Problem	
15.	Easily distracted, concentration wanders	Hyperactivity	
16.	Nervous or clingy in new situations, easily loses confidence	Emotional Symptoms	
17.	Kind to younger children	Prosocial	
18.	Often lies or cheats	Conduct Problem	
19.	Picked on or bullied by other children	Peer Problem	
20.	Often volunteers to help others (parents, teachers, other children)	Prosocial	
21.	Thinks things out before acting (R)	Hyperactivity	
22.	Steals from home, school or elsewhere	Conduct Problem	
23.	Gets along better with adults than with other children	Peer Problem	
24.	Many fears, easily scared	Emotional Symptoms	
25.	Good attention span, sees chores or homework through to the end (R)	Hyperactivity	

Note: R - reverse scoring.

From the answer to these questions from Table A7, 5 outcome variables are constructed. These are –

- **Emotional symptoms:** Items 3, 8, 13, 16, and 24 from the SDQ survey are used to estimate the emotional symptoms score. Each question has a 3-point Likert answering option. By adding the scores of individual questions, the final sub-score is calculated. The variable is continuous; $Y_i^{ess} \in [0,10]$.
- **Conduct problem:** Items 5, 7, 12, 18, and 22 from the SDQ survey are used to estimate the conduct problem score. Each question has a 3-point Likert answering option. The formation of this sub-score is continuous; $Y_i^{cps} \in [0,10]$.

- **Hyperactivity:** Items 2, 10, 15, 21, and 25 are used to estimate the hyperactivity score. The variable is continuous; $Y_i^{hs} \in [0,10]$.
- **Peer problem:** Items 6, 11, 14, 19, and 23 from the SDQ survey are used to estimate peer problem scores. This variable is continuous; $Y_i^{pps} \in [0,10]$.
- **Prosocial:** Items 1, 4, 9, 17, and 20 from the SDQ survey are used to estimate the prosocial score. The variable is continuous; $Y_i^{ps} \in [0,10]$.

A-4.5. Homeschooling time (student)

Students' time investment in homeschooling because of the intervention was measured by asking the following questions to the mother:

a. How much time in various academic activities (i.e., writing, reading, mathematics, etc.) does your child spend per day in minutes?
b. To what extent does your child provide time for academic activities?

Two variables were constructed to underpin the homeschooling time of the students. These are-

- Student's study time: This is the numeric value of daily study time in minutes. The variable is continuous; $Y_i^{ST} \in [0, Y_{max}^{ST}]$.
- Extent of study: This variable is constructed from a 5-point Likert-scale response; 'none' to 'a great deal'. This variable is categorical; $Y_i^{STC} \in \{1..5\}$.

A-4.6. Homeschooling time (Mother or caregiver)

Mother's or caregiver's involvement in children's educational activities was measured by asking the following questions to the mother:

a. How much time do you or the caregiver give to your child in various academic activities (i.e., writing, reading, storytelling, mathematics, etc.) per day in minutes?
b. To what extent do you or the caregiver provide academic support to your child?

Two variables were constructed to underpin the homeschooling time of the students:

- Caregiver's time in home-schooling: This is the numeric value of daily time in minutes given by the mother or primary caregiver in education. The variable is continuous; $Y_i^{PI} \in [0, Y_{max}^{PI}]$.
- Extent of caregiver's time: This variable is constructed from a 5-point Likert-scale response; 'none' to 'a great deal'. This variable is categorical; $Y_i^{PIC} \in \{1..5\}$.

A-5. Social desirability bias

In impact analysis, social desirability bias may arise when the outcome variables rely on survey questions. More specifically, the experimenter demand effect can lead to over or underestimation of the treatment effect. Therefore, during the endline, we surveyed the parents using the short-form Crowne-Marlowe module (Crowne & Marlowe, 1960; Reynolds, 1982) and children using the Children's Social Desirability (CSD-S) scale to measure their social desirability bias (Miller et al., 2015). Questions of these scales are presented in the following Table A8. To explore social desirability bias, we estimated the following interaction model:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 SDB_i + \beta_3 (\mathrm{T} \times \mathrm{SDB})_i + X_i'\theta + \varepsilon_i$$

where SDB_i is a dummy variable that takes a value of 1 if the social desirability score of parents or children is above the median.

Table A8. Social desirability scales – questionnaires

Scale	No	Questions	Desired answer	Answer
Marlowe-Crowne Scale (Reynolds' Form C)	1.	It is sometimes hard for me to go on with my work if I am not encouraged.	False	1) True 2) False
	2.	I sometimes feel resentful when I don't get my way.	False	
	3.	On a few occasions, I have given up doing something because I thought too little of my ability.	False	
	4.	There have been times when I felt like rebelling against people in authority even though I knew they were right.	False	
	5.	No matter who I'm talking to, I'm always a good listener.	True	
	6.	There have been occasions when I took advantage of someone.	False	
	7.	I'm always willing to admit it when I make a mistake.	True	
	8.	I sometimes try to get even rather than forgive and forget.	False	
	9.	I am always courteous, even to people who are disagreeable.	True	
	10.	I have never been irked when people expressed ideas very different from my own.	True	
	11.	There have been times when I was quite jealous of the good fortune of others.	False	
	12.	I am sometimes irritated by people who ask favors of me.	False	
	13.	I have never deliberately said something that hurt someone's feelings.	True	
Children's Social Desirability Short (CSD-S) scale	1.	Have you ever felt like saying unkind things to a person?	No	1) Yes 2) No
	2.	Are you always careful about keeping your clothing neat and your room picked up?	Yes	
	3.	Do you sometimes feel like staying home from school even if you are not sick?	No	
	4.	Do you ever say anything that makes somebody else feel bad?	No	
	5.	Are you always polite, even to people who are not very nice?	Yes	
	6.	Sometimes, do you do things you've been told not to do?	No	
	7.	Do you always listen to your parents?	Yes	
	8.	Do you sometimes wish you could just play around instead of having to go to school?	No	
	9.	Have you ever broken a rule?	No	
	10.	Do you sometimes feel angry when you don't get your way?	No	
	11.	Do you sometimes feel like making fun of other people?	No	
	12.	Do you always do the right things?	Yes	
	13.	Are there sometimes when you don't like to do what your parents tell you?	No	
	14.	Do you sometimes get mad when people don't do what you want them to do?	No	

Notes: This table lists the 13-item short form of the Crowne-Marlowe social desirability scale used for parents and the 14-item Children's Social Desirability (CSD-S) scale used for children.

A-6. Assessment test reliability

We designed our assessment test based on the curriculum designed by the National Curriculum and Textbook Board (NCTB) of Bangladesh. This curriculum has been rigorously tested and modified over the years based on the changing aptitude of the learners. We picked 15 questions from the latest version of the textbooks for the primary grades. These questions are exhibited in Table A4. To verify the coherence of these selected items, we conducted some statistical tests. Firstly, we measured Cronbach's Alpha and McDonald's Omega. Results are exhibited in Table B6. Items of our assessment test have a high level of internal reliability and inter-item correlation. This indicates that assessment test questions are measuring the same underlying construct (literacy/numeracy).

Secondly, we constructed an adjusted assessment test score using a subset of items that appear to perform similarly between treatment and control groups. First, we converted all answers to binary choice (0 = incorrect and 1 = correct). Second, we fitted the answers to all questions to a 2-Parameter Logistic (2PL) Item Response Theory (IRT) model over the treatment and control groups, i.e., a constrained model. Third, we fitted the answers to a specific question to a hybrid 2PL IRT model where parameters can be varied across groups for that item, i.e., an unconstrained model. Fourth, we tested the two models (constrained vs. unconstrained) using the Likelihood-Ratio test to check whether any item shows differential functioning across treatment and control groups.

Figure B6 exhibits the item characteristics curves for all 15 questions from the endline assessment tests and Table B7 exhibits LR test statistics. Based on the LR test, it is evident that few items exhibit differential item functioning. Finally, we re-estimated the assessment test score by excluding these items. These adjusted treatment effects are presented in Table B8, which indicates that the learning outcomes are improved after Differential Item Functioning (DIF) adjustment. Overall, our assessment test is stable, reliable, and coherent and the treatment was effective based on the DIF-adjusted test score.

Appendix B: Additional Figures and Tables

Figure B1. IVR flow diagram of T1: Standard treatment group.

04445 --- --- Dial this number. Your call will be hanged up and called back.

Greeting message and welcome to IVR-based mobile school. Module option will be given.

Press 1 for Literacy
Press 2 for Numeracy

Lesson option of the selected module will be given.

There are 30 lessons in Literacy and Numeracy. Press any lesson number to listen.

Listen the selected lesson. A quiz will be played after each lesson.

Press 1, 2 or 3 to give the answer of the quiz.

Option to listen more lesson will be given.

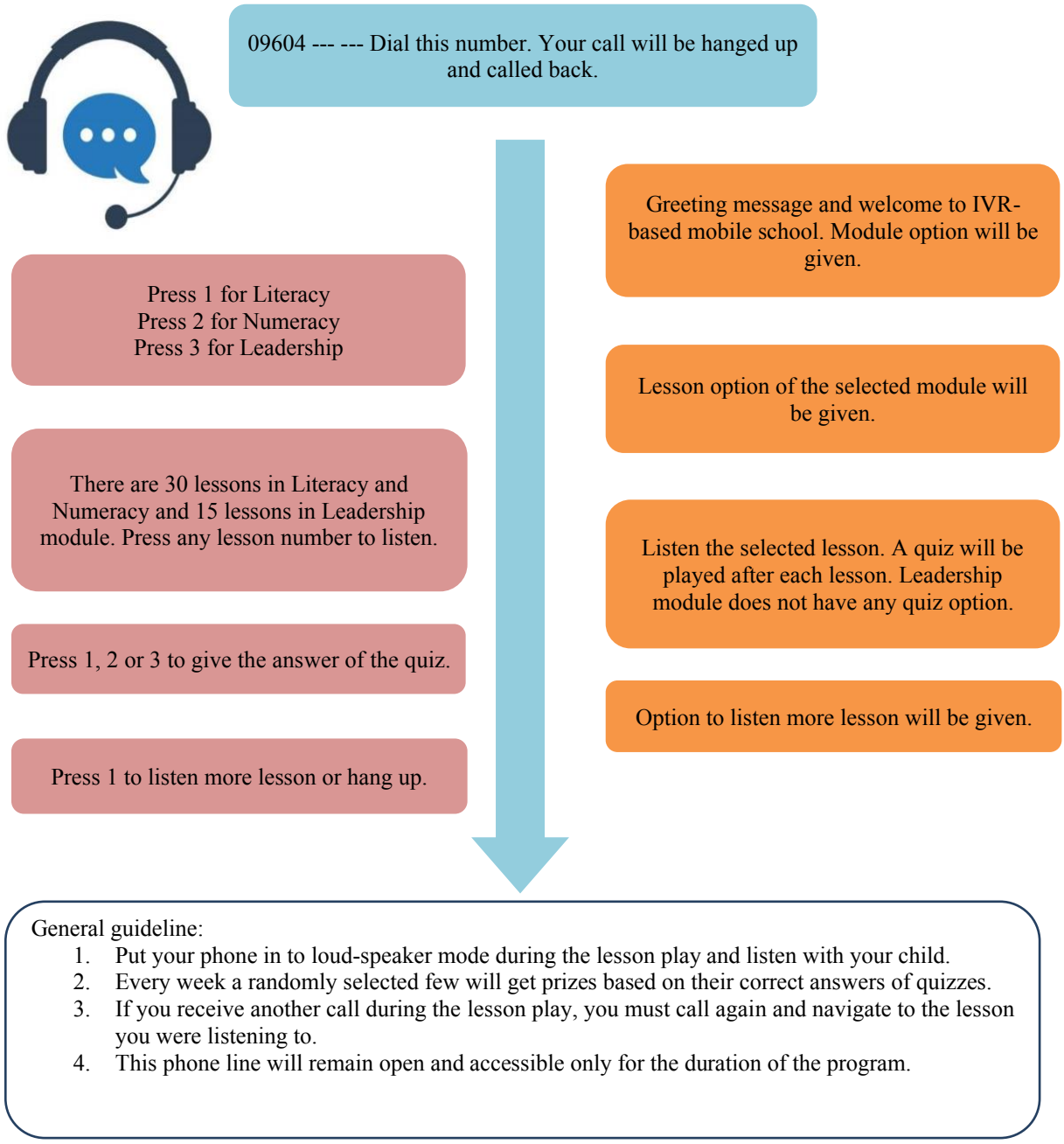
Press 1 to listen more lesson or hang up.

General guideline:

1. Put your phone in to loud-speaker mode during the lesson play and listen with your child.
2. Every week a randomly selected few will get prizes based on their correct answers of quizzes.
3. If you receive another call during the lesson play, you must call again and navigate to the lesson you were listening to.
4. This phone line will remain open and accessible only for the duration of the program.

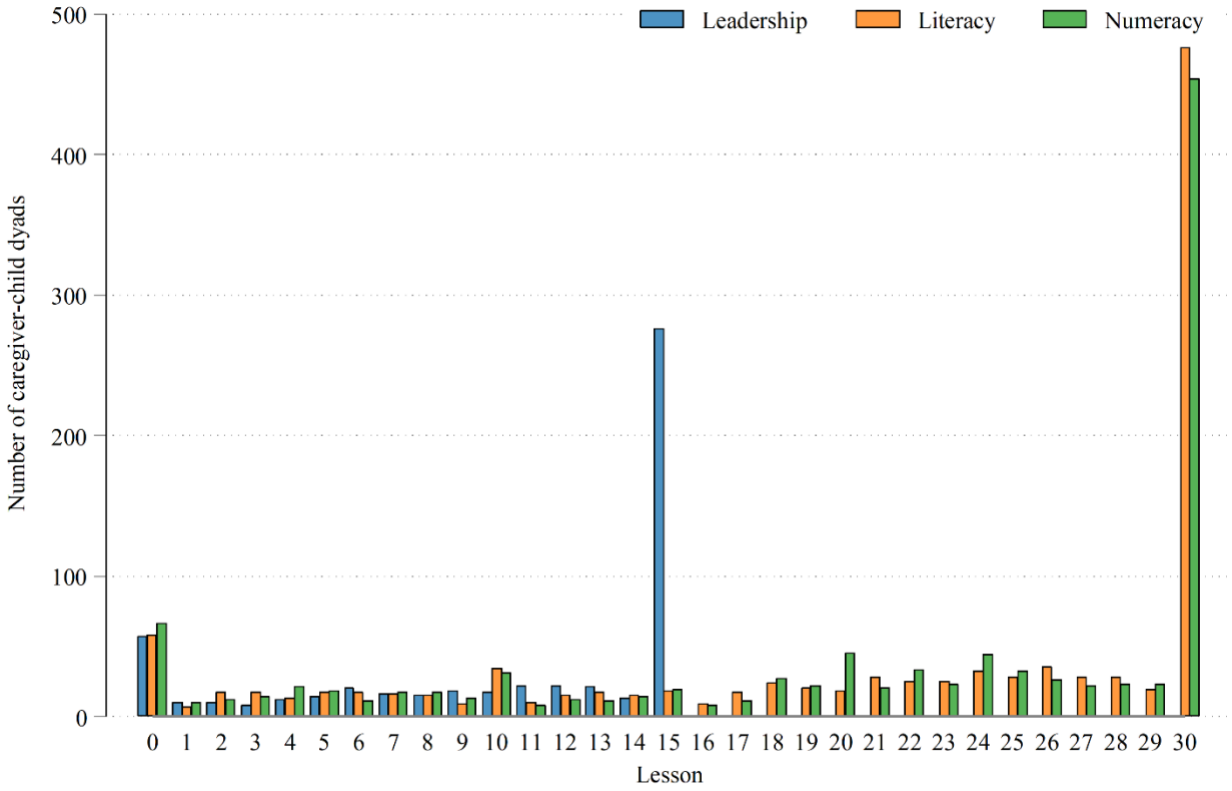
Note: This figure illustrates the IVR journey of the participating caregiver-child pairs in the Standard treatment group.

Figure B2. IVR flow diagram of T1: Extended treatment group.



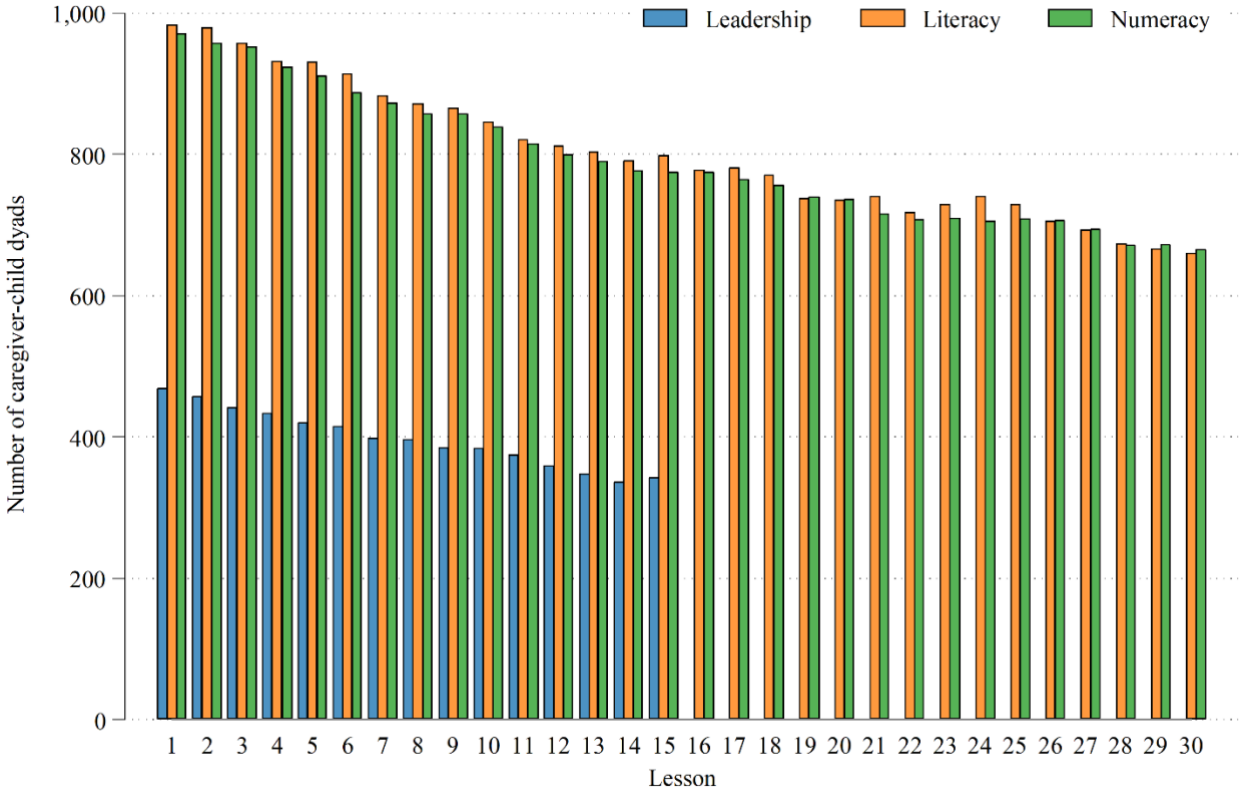
Note: This figure illustrates the IVR journey of the participating caregiver-child pairs in the Extended treatment group.

Figure B3. Number of lessons listened by caregiver-child dyads, by module.



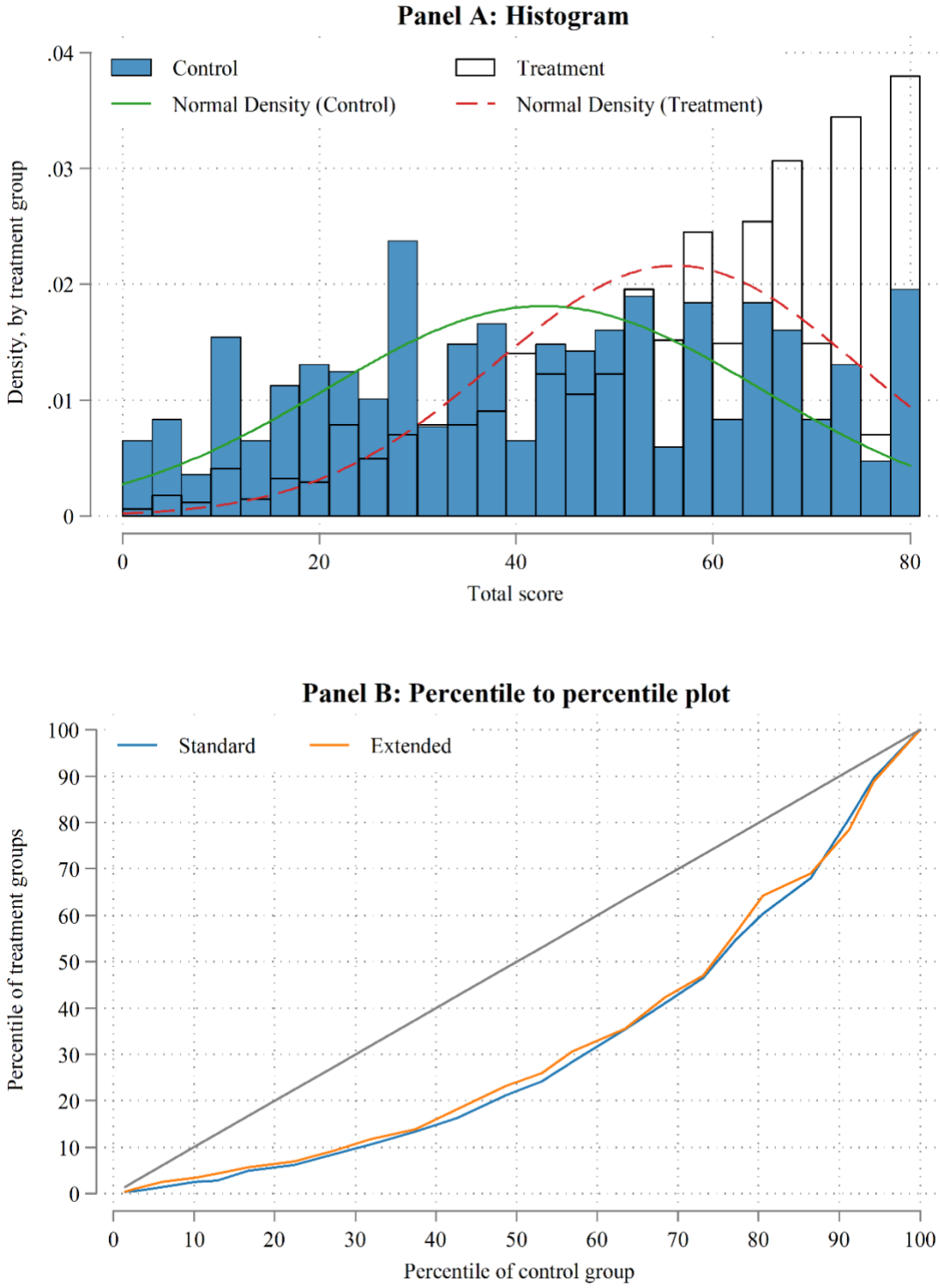
Note: This figure shows the number of lessons completed by the caregiver-child dyads in different modules. In Literacy, Numeracy and Leadership, 58, 66 and 57 dyads did not listen to any of the lessons, respectively. On the contrary, 476, 454 and 276 dyads received all the lessons for Literacy, Numeracy and Leadership, respectively. The Leadership module was offered only to the Extended group and there were 15 lessons. In our intervention, a total of 1,182 caregiver-child dyads received the treatment. However, a total of 1107 dyads fill-up the lesson attendance sheet. The rest of the 75 dyads did not fill-up the form and is therefore excluded from these charts.

Figure B4. Attendance of caregiver-child dyads, by lesson.



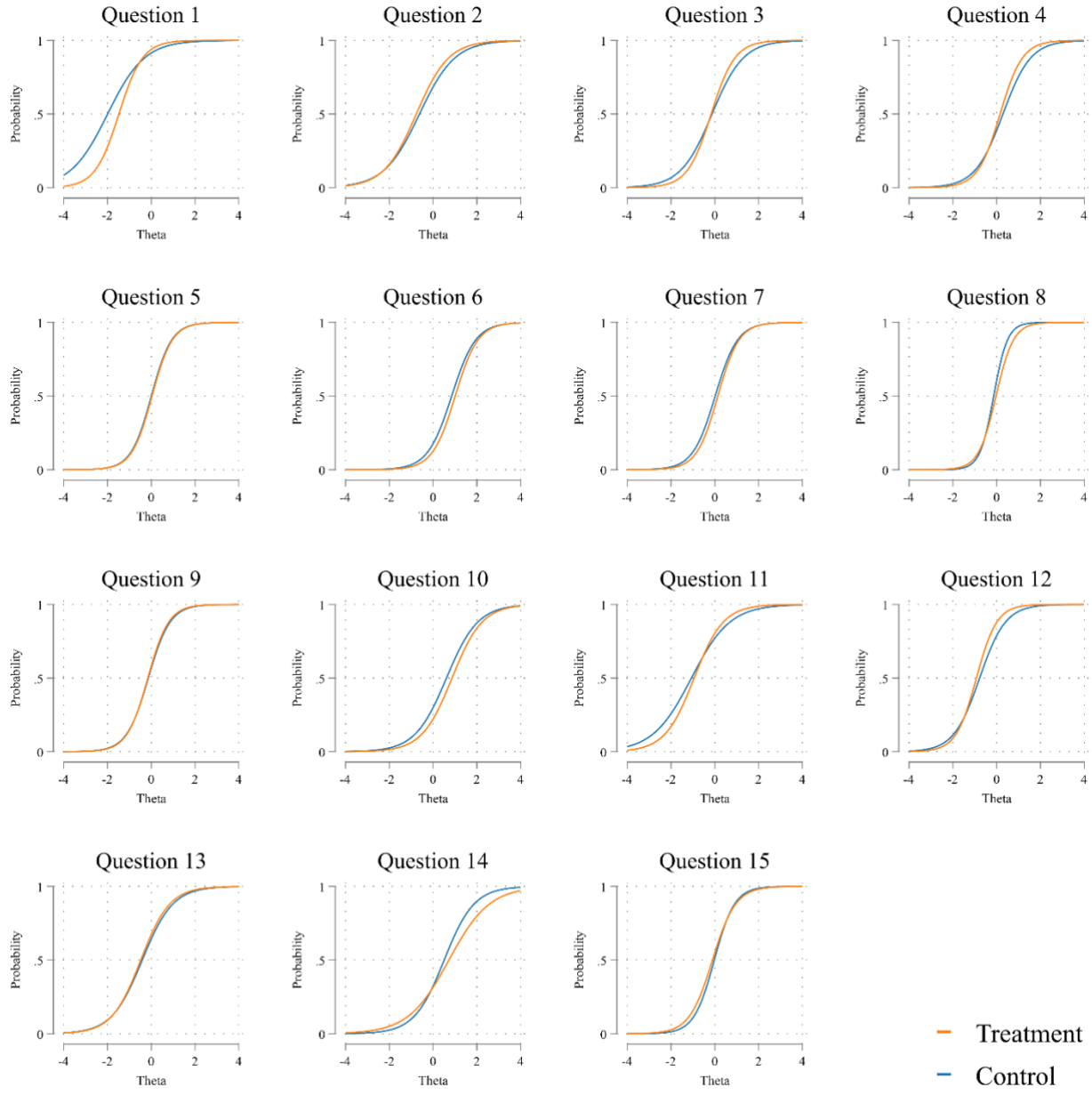
Note: This figure shows the attendance of caregiver-child dyads in different lessons of different modules. In the Leadership module, there were only 15 lessons. In our intervention, a total of 1,182 caregiver-child dyads received the treatment. However, a total of 1107 dyads fill-up the lesson attendance sheet. The rest of the 75 dyads did not fill up the form and is therefore excluded from this figure.

Figure B5. Assessment-test score distribution, by treatment.



Note: This figure exhibits a comparison between test-score distributions. In Panel A, histograms of treatment and control groups are presented. Test score of the treatment group is left-skewed, which means treated children scored higher than untreated children. In Panel B, percentiles of treatment groups and control groups are exhibited. This figure indicates that the 30th percentile of the treatment group distribution corresponds to the 60th percentile of the control group distribution.

Figure B6. Differential Item Functioning (DIF).



Note: This figure exhibits the probability of answering each question correctly, by estimated ability (theta), and by treatment (pooled).

Supplementary Tables

Table B1. Survey attrition rate.

Particulars	(1)		(2)		(3)		(4)	
	T1: Standard		T2: Extended		T3: Control		Total	
	N	%	N	%	N	%	N	%
Not attrited	567	95.13%	562	95.90%	561	96.56%	1690	95.86%
Attrited	29	4.87%	24	4.10%	20	3.44%	73	4.14%
Total	596	100%	586	100%	581	100%	1763	100%

Note: This table reports the frequency of attrition at the endline survey. All 1763 children were approached for the endline survey and assessment. A total of 1690 child and mother dyads completed both child assessment and parent survey. Pearson's Chi-squared test: χ^2 = 1.5063, P = 0.47.

Table B2. Attrition, by treatment.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	T1 & T3	T1 & T3	T2 & T3	T2 & T3	T1, T2 & T3	T1, T2 & T3
T1: Standard treatment dummy	0.01	0.06	-	-	-	-
	(0.01)	(0.17)				
T2: Extended treatment dummy	-	-	0.01	-0.24	-	-
			(0.01)	(0.20)		
Any treatment Dummy	-	-	-	-	0.01	-0.09
					(0.01)	(0.17)
Constant	0.03***	0.30**	0.03***	0.30**	0.03***	0.30**
	(0.01)	(0.14)	(0.01)	(0.14)	(0.01)	(0.14)
Controls	No	Yes	No	Yes	No	Yes
Interaction terms (treatment dummy × controls)	No	Yes	No	Yes	No	Yes
Observations	1,177	1,177	1,167	1,167	1,763	1,763
R-squared	0.00	0.13	0.00	0.11	0.00	0.11
Joint F-test *p-value* on characteristics	-	0.01	-	0.01	-	0.01
Joint F-test *p-value* on interactions	-	0.67	-	0.84	-	0.85

Note: All columns present estimates using OLS, where the dependent variable is a dummy variable for attrition (1 if the parent-child dyad did not participate in the endline survey and assessment). T1 – Standard treatment; T2 – Extended treatment; and T3 – Control. The sample in columns 1 & 2 is parent-child dyads in the T1 & T3, the sample in columns 3 & 4 is parent-child dyads in the T2 & T3, and the sample in columns 5 & 6 is the full sample. Control variables are – children's age, gender, baseline literacy score, baseline numeracy score, access to private tuition, parents' education in years, family income, religion, access to TV & smartphone, homestead size and the number of members in the household. Robust standard errors clustered at the village level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B3. Percentage of children who answered correctly, by question and treatment.

Module	Question no	Type	(1)	(2)	(3)	(4)	(5)
			Marks assigned	Standard	Extended	Control	Total sample
Bangla literacy	1	Letter identification/ Reading	5	0.94 (0.01)	0.96 (0.01)	0.87 (0.02)	0.92 (0.01)
	2	Vocabulary	5	0.80 (0.02)	0.84 (0.01)	0.64 (0.03)	0.76 (0.01)
	3	Vocabulary/ Spelling	5	0.74 (0.02)	0.73 (0.02)	0.53 (0.03)	0.67 (0.01)
	4	Vocabulary/ Spelling	5	0.65 (0.03)	0.66 (0.02)	0.42 (0.03)	0.57 (0.01)
	5	Letter identification/ Reading	6	0.70 (0.03)	0.70 (0.02)	0.50 (0.03)	0.63 (0.01)
	6	Answering verbally to an English question	6	0.41 (0.03)	0.38 (0.03)	0.27 (0.02)	0.35 (0.01)
English literacy	7	Vocabulary	4	0.67 (0.02)	0.66 (0.02)	0.49 (0.03)	0.61 (0.01)
	8	Vocabulary	4	0.73 (0.02)	0.74 (0.02)	0.54 (0.04)	0.67 (0.01)
	9	Vocabulary	4	0.73 (0.03)	0.76 (0.02)	0.54 (0.03)	0.68 (0.01)
	10	Writing/ Spelling	6	0.45 (0.03)	0.45 (0.03)	0.35 (0.03)	0.41 (0.01)
Numeracy	11	Basic numbers	6	0.87 (0.02)	0.85 (0.02)	0.73 (0.02)	0.82 (0.01)
	12	Addition/ Sorting	6	0.89 (0.01)	0.91 (0.01)	0.71 (0.02)	0.84 (0.01)
	13	Addition/ Subtraction	6	0.78 (0.02)	0.79 (0.02)	0.60 (0.03)	0.72 (0.01)
	14	Number conversion	6	0.49 (0.03)	0.49 (0.03)	0.37 (0.03)	0.45 (0.01)
	15	Addition/ Multiplication	6	0.73 (0.03)	0.72 (0.02)	0.50 (0.03)	0.65 (0.01)
All		Number of questions answered correctly	-	10.59 (0.23)	10.65 (0.20)	8.06 (0.31)	9.77 (0.09)

Note: This table exhibits the percentage of children answered correctly in each question of the endline assessment test by treatment groups in columns 2-4. Column 5 indicates the percentage for the full sample. Robust standard errors clustered at the village level are in parentheses. All these differences are statistically significant.

Table B4. Parents' social desirability bias analysis.

Variables	T1: Standard			T2: Extended		
	(1) Above median	(2) Below median	(3) Interaction	(4) Above median	(5) Below median	(6) Interaction
Leadership	0.08	0.15	-0.08	0.14	0.13	0.01
	(0.09)	(0.14)	(0.13)	(0.09)	(0.12)	(0.13)
Communication	0.03	0.18	-0.14	-0.00	0.06	-0.08
	(0.10)	(0.16)	(0.16)	(0.12)	(0.12)	(0.14)
Planning	-0.01	0.15	-0.15	-0.02	0.14	-0.17
	(0.12)	(0.16)	(0.16)	(0.13)	(0.13)	(0.15)
Emotional symptoms	-0.14	-0.19**	0.05	-0.11	-0.16	0.04
	(0.10)	(0.09)	(0.13)	(0.09)	(0.10)	(0.13)
Conduct problem	-0.10	-0.32***	0.24*	-0.13**	-0.08	-0.02
	(0.07)	(0.12)	(0.13)	(0.06)	(0.11)	(0.11)
Hyperactivity	-0.02	-0.40***	0.40***	0.05	-0.04	0.12
	(0.08)	(0.10)	(0.13)	(0.07)	(0.12)	(0.12)
Peer problem	-0.14	-0.18*	0.06	-0.01	-0.17	0.18
	(0.09)	(0.10)	(0.14)	(0.08)	(0.12)	(0.14)
Prosocial	0.11	0.44***	-0.33***	0.11	0.14	-0.04
	(0.08)	(0.11)	(0.11)	(0.09)	(0.09)	(0.11)
Student's study time	0.07	0.18*	-0.08	0.10	0.02	0.06
	(0.12)	(0.10)	(0.15)	(0.11)	(0.11)	(0.13)
Caregiver's time in homeschooling	0.15	0.30**	-0.15	0.05	0.10	-0.07
	(0.10)	(0.12)	(0.14)	(0.10)	(0.11)	(0.13)

Note: This table exhibits treatment effects adjusted for the social desirability bias of the parents. Only those outcome variables that are dependent on parents' survey answers are included in this analysis. We estimated the social desirability bias of the parents using the 13-item short form of the Crowne-Marlowe social desirability scale. Above median – dummy variable for above-median value of social desirability bias scale; interaction – interaction term between SDB and treatment dummy. Coefficients are estimated with OLS. The dependent variable for each regression is listed in the row heading. The same list of control variables is used as before. Children's grade fixed effects are included in all regressions. Robust standard errors clustered at the village level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B5. Child's social desirability bias analysis.

Variables	T1: Standard			T2: Extended		
	(1) Above median	(2) Below Median	(3) Interaction	(4) Above median	(5) Below Median	(6) Interaction
Overall impulsivity	-0.18*	-0.07	-0.13	-0.11	0.13	-0.26**
	(0.09)	(0.12)	(0.11)	(0.10)	(0.13)	(0.12)
Schoolwork impulsivity	-0.27***	-0.10	-0.19*	-0.19*	0.08	-0.30**
	(0.10)	(0.12)	(0.11)	(0.10)	(0.12)	(0.13)
Interpersonal impulsivity	-0.05	-0.03	-0.04	-0.02	0.14	-0.17
	(0.08)	(0.12)	(0.11)	(0.09)	(0.11)	(0.12)
Grit	0.15	0.07	0.07	0.14	0.07	0.05
	(0.11)	(0.11)	(0.12)	(0.12)	(0.10)	(0.12)
Growth mindset	0.19	0.11	0.10	0.25**	0.20	0.05
	(0.13)	(0.13)	(0.13)	(0.12)	(0.12)	(0.14)
Affective empathy (Contagion)	-0.08	-0.13	0.04	-0.05	-0.18**	0.12
	(0.11)	(0.08)	(0.14)	(0.12)	(0.08)	(0.13)
Cognitive empathy (Understanding)	-0.15	-0.13	-0.00	-0.11	-0.12	-0.00
	(0.10)	(0.10)	(0.14)	(0.09)	(0.09)	(0.12)
Prosocial motivation (Support)	-0.07	-0.02	-0.03	-0.05	-0.06	-0.00
	(0.12)	(0.09)	(0.14)	(0.11)	(0.07)	(0.13)

Note: This table exhibits treatment effects adjusted for the social desirability bias of the children. Only those outcome variables that are dependent on children's survey answers are included in this analysis. We estimated social desirability bias using the Children's Social Desirability (CSD-S) scale. Above median – dummy variable for above-median value of social desirability bias scale; interaction – interaction term between SDB and treatment dummy. Coefficients are estimated with OLS. The dependent variable for each regression is listed in the row heading. The same list of control variables is used as before. Children's grade fixed effects are used in all regressions. Robust standard errors clustered at the village level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B6. Reliability of endline assessment test questions.

Subject	Alpha	Omega	Correlation	No of questions	No of students
All questions	0.86	0.86	0.28	15	1690
Literacy	0.82	0.82	0.31	10	1690
Numeracy	0.70	0.70	0.32	5	1690

Note: This table shows Cronbach's Alpha, McDonald's omega, and the inter-item correlation for the assessment instruments used in the endline.

Table B7. Differential Item Functioning (DIF) analysis.

Subject	Question number	Chi-square	P-value
Literacy	1	4.76	0.09
	2	3.73	0.15
	3	5.27	0.07
	4	5.34	0.07
	5	0.17	0.92
	6	4.85	0.09
	7	2.75	0.25
	8	6.99	0.03
	9	0.24	0.89
	10	7.77	0.02
Numeracy	11	3.64	0.16
	12	11.79	0.00
	13	0.99	0.61
	14	5.95	0.05
	15	1.76	0.42

Note: This table exhibits the likelihood ratio (LR) test statistics test for differential item functioning (DIF) of various questions of the assessment tests. The null hypothesis is no DIF in respective items across treatment and control groups. All questions are converted to dichotomous, 0 – wrong answer, 1 – correct answer.

Table B8. Differential Item Functioning (DIF) adjusted learning outcome.

Variable	T1: Standard		T2: Extended	
	Coefficient on treatment dummy	t-test *P-value*	Coefficient on treatment dummy	t-test *P-value*
Total score	0.60***	0.00	0.62***	0.00
	(0.08)		(0.07)	
Literacy	0.56***	0.00	0.59***	0.00
	(0.08)		(0.07)	
Numeracy	0.49***	0.00	0.50***	0.00
	(0.07)		(0.06)	

Note: This table exhibits DIF-adjusted learning outcomes. It was evident that out of fifteen questions, there was DIF in four questions. We re-estimate our learning outcomes excluding these questions and then estimate the standardized treatment effects. Coefficients are estimated with OLS. The dependent variable for each regression is listed in the first row. The same list of control variables is used as before. Children's grade fixed effects are used in all regressions. Robust standard errors clustered at the village level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.