

Supporting Early Childhood Development Remotely: Experimental Evidence from SMS Messages*

Juan Manuel Hernandez-Agramonte (IPA)[†] Olga Namen (UC Santa Barbara)[‡]
Emma Naslund-Hadley (IDB)[§] Maria Loreto Biehl (IDB) [¶]

May 2, 2022

Abstract

Using a randomized field experiment in Costa Rica, we estimate the effect of providing parents of preschool students with a text message intervention containing information and activities to engage them with their children's learning process at home. After 15 weeks of intervention, the cognitive skills of children whose parents were assigned to the program was 0.11-0.12 standard deviations higher than the control group. We find suggestive evidence that the effect was driven by an increase in parent involvement through the proposed activities from the text message campaign.

Keywords: Remote Education, Early Childhood Development, Text Messages, Parenting, COVID-19, Parental Networks, Parental School Involvement.

JEL Codes: C93, I21, J13, O15.

*We thank Kelly Bedard, Steve Glazerman, Heather Royer, Dick Startz, Gonzalo Vazquez-Bare and participants at the UCSB Applied Micro Lunch, the UCSB Econ 290 Research Group and the "Evidence Day Summit - National Education Week" organized by the Ministry of Education of Ghana and IPA for their helpful comments and suggestions. We gratefully acknowledge support from Melania Brenes and Guisselle Alpizar at the Ministry of Public Education of Costa Rica and thank Carina Castro for her valuable input on the design of the text message campaign. Joaquin Armas, Carmen Mira-Alonso, Kelly Montaña, Rayssa Ruiz, Carolina Saavedra, and Carlos Urrutia provided outstanding research assistance. Innovations for Poverty Action assisted with the project management, field work, and Institutional Review Board (IRB) review. This research was funded by the Inter-American Development Bank (IDB). AEA RCT Registration Number: AEARCTR-0006955. IPA IRB Protocol: 15517. The opinions expressed are those of the authors and do not necessarily reflect the views of the IDB, its board of directors, or the countries they represent.

[†]jmhernandez@poverty-action.org

[‡]onamen@ucsb.edu

[§]emman@iadb.org

[¶]loretob@iadb.org

1 Introduction

Early childhood is a critical life period for the development of abilities and skills that are decisive for future outcomes (Heckman, 2006). Research documents large benefits to individuals and society from educational investments during early childhood, especially for children living in disadvantaged conditions (Elango et al., 2015). In emergency contexts, families face very exceptional constraints to provide sufficient stimulation during these early years (Bouchane et al., 2019). Recently, the COVID-19 pandemic forced education worldwide to shift from classroom-based instruction to entirely remote learning strategies, in which education occurs in the student's home. This situation imposed enormous challenges for the educational system, particularly in developing countries, where connectivity and household resources are limited. Educating preschool students poses an additional challenge, as they require active support from their parents to access and use distance education resources. Preschool students in low-income households may be particularly adversely affected by school closures because their parents often have fewer resources and skills and face greater economic and psychological impacts during emergencies.

This paper evaluates the impacts of a text message intervention implemented in Costa Rica designed to support preschool students' learning at home during the lockdown imposed by the COVID-19 emergency. We conduct a two-stage experiment to estimate the direct and spillover effects of the intervention on children's early cognitive outcomes. The program targeted parents of four- and five-year-old preschool students from public schools across the country. The text messages were designed to increase parental engagement with their children's education by providing simple learning activities. In addition to the educational content, the program sought to create a more conducive environment for learning at home by addressing various dimensions of parenting, such as parenting style, time management, and healthy habits.

A distinctive feature of this intervention is that it was implemented in a setting where traditional in-person instruction was disrupted, as schools in Costa Rica were closed in response to the pandemic (Näslund-Hadley et al., 2020). Most communication between the education providers and parents was done through phone calls and text messages. Teachers, in particular, set up WhatsApp groups for their classrooms, through which they shared information and resources with parents. This virtual environment provides an opportunity for parents to use these channels of communica-

tion to share information about their children’s education and interact with each other around these issues. We design our randomized experiment to capture spillover effects of the intervention within the groups of parents that share a preschool teacher (parental network). The experimental sample includes 691 parental networks and 4,496 students. In the first stage, we randomly assign networks to either a treatment group or a pure control group. In the second stage, we randomly assign half of the parents in each treatment network to either receive the text messaging intervention (treated group) or to not receive the intervention (untreated group). Comparing the outcomes across these experimental groups of parents—treated, untreated, and pure control—allows us to estimate the direct effect of the text messaging campaign as well as the spillover effects within the networks.

As school closures imposed limitations on face-to-face assessments to measure children’s cognitive outcomes, we develop an innovative remote learning assessment by adapting the Early Grade Mathematical Assessment (EGMA) and the Measuring Early Learning Quality and Outcomes (MELQO) tests. The assessment was administered over the phone, to overcome potential Internet connectivity restrictions. Since outcomes were measured about a week after the text messaging program ended, the effects should be interpreted as short-term. We also collect baseline and follow-up data on parents and household characteristics through online surveys, which allow us to measure changes in parental investments and home environment.

We find that the text message program had positive effects. Results show that cognitive skills of students whose parents were assigned to receive the text message campaign are 0.11– 0.12 standard deviation higher than the control group. The effect corresponds mainly to an improvement in early numeracy skills. Effects are larger in networks that had lower baseline parent-teacher interaction suggesting that messages are more effective in less connected households where access to this information would have been hard to obtain otherwise. These short-term impacts are substantial, taking into account that the intervention lasted just 15 weeks. Consistent with the effects on students’ cognitive skills, we find suggestive evidence of higher parent involvement through the proposed activities from the text messaging campaign. These results are in line with existing evidence that inexpensive behavioral interventions can empower parents to improve children’s outcomes (Bergman, 2019). We don’t find strong evidence of changes in cognitive skills for students whose parents were part of treatment networks but did not receive the messages themselves. Given the evidence of direct treatment effects, this result is somewhat surprising as priori we expected a

stronger role of the networks to compensate for the lack of preschool instruction. The absence of spillover effects is robust in most specifications.

Our paper provides two main contributions to the literature studying parental engagement in early childhood (Hurwitz et al., 2015; Doss et al., 2019; Mayer et al., 2019; York et al., 2019; Barrera et al., 2020; Cortes et al., 2021). First, our experiment is designed to measure potential spillover effects in addition to evaluating direct impacts of the intervention. This allow us to examine the dynamics of parental behavior and children’s cognitive development within their existing networks. Second, we show that text-message interventions can be effective even in a context where preschool centers are closed and children are receiving services remotely at home. These findings have important implications for the provision of early childhood education in emergencies and conflict settings which, according to recent research, is an area that has not been prioritized or funded enough compared to other humanitarian aids (Ponguta et al., 2022). Text message interventions offer a low-cost alternative for addressing school disruptions when children don’t have access to in-person care and high-technology solutions are less scalable given the limited access to the Internet or digital devices.

The paper is organized as follows: Section 2 provides background of the setting and the program. Section 3 describes the experimental design. Section 4 outlines the data sources. Section 5 illustrates the empirical strategy. Section 6 discusses the validity checks. Section 7 reports the results. Section 8 discusses the results and explores potential mechanisms. Finally, Section 9 presents our conclusion.

2 Intervention

2.1 Institutional background

According to the OECD, Costa Rica has sharply expanded access to quality preschool education by increasing participation in 4-year-old level from 7% to 63% and in 5-year-old level from 83% to 90% (OECD, 2017). In fact, Costa Rica was one of the first countries in Latin America to adopt COVID-19 pandemic preventive measures, closing its schools on March 12, 2020. The Ministry of Public Education (MEP) launched the national remote learning program *Aprendo en Casa (AeC)* to continue with the school year, that, as most school calendars in Latin America, it

had just recently began. The program combined printed materials and technology-based solutions to address heterogeneous household access to telecommunications, aiming to reach the largest share of students. Educational content was uploaded to MEP's website and adapted to be broadcast in TV and radio.

Teachers were instructed to communicate with their students' families and provide additional materials and guidance, typically through phone calls and WhatsApp groups, where Internet was available. According to our baseline data, 98% of teachers established a communication channel with families and students, 68% were able to communicate with all of their families in the class, and an additional 25% were able to communicate with more than half of the families in the class. The main purpose of this communication was to supervise how parents and students progressed with their study materials, to solve questions, and to support parents in providing learning guidance to their children. Teachers reported that most parents have initiated communication with them and that most parents are following the guidance they receive to practice at home.

MEP was quick to provide an alternative to school-based education, but the strategy faced several challenges. For preschool students, adults need to help their children access the strategy materials and guide them through the learning process, but few parents have the training or experience to guide their children's formal education at home. MEP provided weekly support to parents; 84% of parents reported that someone at MEP had communicated with them to discuss AeC and motivate parents to support their children's learning. This interaction was mainly through WhatsApp messages, video call software, and phone calls. Parents found that supporting their children's education at home was challenging, with 56% reporting that they needed more help or support from other family members. Their main requests were receiving more learning activities to implement at home and more directions on how to implement those activities. Adding to the challenges of implementing the remote educational process was the fact that the pandemic had severe effects on the home environment and parents mental health toll. Our survey data show that 49% of parents presented at least one symptom of frequent stress. These dimensions are out of the programmatic scope of the ministries of education of the region, but many of them have acknowledged the need to include components that address home environments in their remote education programs (Näslund-Hadley et al., 2020).

2.2 Text message program

Together with the MEP, we designed a text message program to support the education process at home during the pandemic. The campaign was implemented over a period of 15 weeks (August 24–November 30, 2020). Figure 1 provides a timeline of the project. The intervention focused on parents with four- or five-year-old children that were enrolled in public preschools.

Parents in the treatment group received a series of 3–4 weekly messages, up to a total of 54 messages. Text messages were prescheduled to be sent on particular days and times of the week between 3 pm and 5 pm using a bulk messaging platform. The messages aimed to increase child learning by providing parents with examples of simple numeracy and literacy activities to do with their children.¹ The activities were based on MEP’s preschool curriculum and designed to be implemented at home with no additional support materials.² The messages also included a weekly motivational message that prompted parents to implement the activities.³ The campaign complemented the educational and motivational messages with advice on positive parenting, time management, and healthy habits intended to help parents create a conducive home environment for learning.⁴

Parents in the untreated and control groups received a shorter information campaign with 8 simple messages reminding parents about channels for accessing AeC. All three groups received regular virtual services from AeC.

¹Numeracy skills included counting, adding, subtracting, comparing numbers, and sequencing. Literacy skills included oral comprehension, expressive vocabulary, and breaking words into syllables.

²For example, “Let’s play counting! Ask your child to count objects like rocks, spoons or beans. Encourage him to reach a higher number each time. Practice daily!” “Let’s do addition! Ask your child: If you have four objects and I give you two more, how many objects will you have? Practice every day with different quantities.”

³The messages used behavioral tools to address common parent biases regarding early education. They included information on the returns to early childhood development to address inaccurate beliefs about the importance of early ages learning. Messages with positive affirmations of parents’ ability to ensure children learning were also included. Some of these messages were combined with a loss aversion framing and positive reinforcement. For example, “Did you enjoy when your child showed you how he counted objects? That and more they can learn everyday with your love and support. Play this week’s games and have fun!”

⁴These messages included tips for planning weekly activities with the child, managing stress and having effective parent–children interactions. In addition, messages recommended doing physical activity, breathing exercises, singing, playing and avoiding of long hours in front of screens

3 Experimental design

We worked with the preschool education unit of MEP to get access to teachers and recruit parents from their classes. We defined a network as a group of parents that share the same preschool teacher. During the pandemic, teachers used different means to communicate with parents and coordinate the implementation of the remote education strategy. One of the most common channels was WhatsApp groups, which enabled teachers to interact with parents using a variety of formats, including text, audio, images, and videos. We asked teachers to share with parents a link with a short introduction to the project and an online survey. The survey invited parents of preschool children to join the study and also collected baseline information. Our experimental design required networks in which at least two parents consented to participate.

We used a two-stage randomized experiment to quantify the direct and spillover short-term effects of providing the text message campaign. First, we conducted a group-level randomization at the network level to divide the sample into pure control networks and treatment networks. We stratified by province where the child center is located, network size, and the level of interaction between parents and teacher within the network. Figure 2 shows the distribution of network size. We used an indicator of big-sized networks constructed as whether the number of parents within the network was higher than the median. For the level of interaction we used an indicator of high interaction constructed as whether the teacher reported that all or almost all parents in her network have communicated with her. Since there are 7 provinces, we have 28 strata in total.

Second, we conducted an individual-level randomization within treatment networks to randomly assign half of the parents to either receive the text message intervention (treated units) or not receive the intervention (untreated units). This experimental design allows us to estimate spillover effects as the difference in outcomes between the untreated parents in treatment networks and parents in pure control networks. We estimate direct effects as the difference in outcomes between treated parents in treatment networks and parents in pure control networks.⁵

As shown in Figure 3, our experimental sample consists of 4,496 parents. Of the 691 networks in the study, 338 networks were randomly assigned to the pure control group and 353 networks were randomly assigned to the treatment group. Within the treatment group, 1,072 parents were

⁵Further details on partial population designs, see Moffit (2001); Duflo and Saez (2003); Baird et al. (2018); Vazquez-Bare (2022).

randomly assigned to receive the text messaging intervention (treated group), and 1,250 were not assigned to receive it (untreated group).⁶

4 Data

For our analysis we rely mainly on two sources of information: (i) online surveys administered to teachers and parents at baseline to recruit participants and capture pretreatment characteristics and (ii) a phone-based learning assessment to measure children’s cognitive skills, including early numeracy and literacy skills. The endline child learning assessment was collected in December after the text message program was complete. See Figure 1 for a timeline of the project.

(i) Parent and teacher baseline survey. We collected baseline information on teachers and parents by leveraging the existing communication channels between the MEP, preschool teachers, and students’ parents. Teachers shared with parents an online survey that requested consent from parents to join the study and collected information on pretreatment characteristics, including household demographics, socioeconomic information, baseline information on the home learning environment and access to play-materials, level of stress from children and parents and use of violent discipline practices. See Appendix B.1 for more information about the variables collected.

(ii) Child learning assessment. Our main outcomes of interest are children’s cognitive skills. During the implementation of this evaluation, Costa Rica was under a strict lockdown that prevented face-to-face data collection. Therefore, we adapted the Early Grade Mathematical Assessment (EGMA) and the Measuring Early Learning Quality and Outcomes (MELQO) tests into a phone-based learning assessment. The assessment was administered over the phone to overcome Internet connectivity restrictions. The phone-based assessment consisted of multiple numeracy questions, including questions on spatial reasoning, oral counting, comparisons, addition, subtraction and sequences. The assessment also measured a set of early literacy skills, including syllabication, oral comprehension and expressive vocabulary. We compute a composite score of children’s cognitive skills as well as numeracy score and a literacy score. Each score was standardized so that the

⁶Our power calculations indicate that with this sample size, assuming an intra-cluster correlation of 0.1, a probability of error type I of 0.05, and a power of 0.8, the minimum detectable effect is 0.12 for the child’s standardized cognitive score.

pure control group has a mean of zero and a standard deviation of one. The test was administered by enumerators who called the parents to schedule a time to test their child. During the call, parents were provided support to set up the phone call in speaker mode. We incorporated a series of procedures to increase the reliability of the measures collected. We instructed parents to prepare a place for their child to take the test without distractions. Parents were told that the test was low-stakes, in order to minimize their interference, and were instructed not to help the child during the test. We used reminders in different parts of the assessment asking parents not to interrupt the call. In terms of validity, the Cronbach's alpha coefficient across the 29-item scale is 0.8217, suggesting that items have high internal consistency to measure cognitive ability. In Appendix B.2 we describe this assessment in detail and use Item Response Theory to provide validity evidence for the use of this instrument tool.

5 Empirical strategy

To evaluate the direct and spillover effects of the text message program, we estimate the following ordinary least squares (OLS) regression:

$$Y_{ig} = \alpha_s + \beta_1 D_{ig} + \beta_2 (1 - D_{ig}) \times T_g + X_{ig} + \epsilon_{ig}, \quad (1)$$

where Y_{ig} is the outcome for each child i in network g ; D_{ig} is an indicator of whether the household was assigned to receive the text message campaign (treated group); and $(1 - D_{ig}) \times T_g$ is an interaction term that indicates households that belong to a treatment network but where themselves not assigned to receive the text messages (untreated group). Households in pure control networks are the omitted category; X_{ig} is a matrix of baseline characteristics, including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline.⁷ The α_s are the randomization strata

⁷See Appendix B.1 for more information about the construction of these variables.

fixed effects and ϵ_{ig} is the error term. We cluster standard errors at the network level, allowing for correlation between disturbances of students under the same teacher. Parameter β_1 is interpreted as the direct effect of being selected to receive the text message campaign. Parameter β_2 corresponds to the spillover effect of being part of a parental network in which other parents were selected to receive the text message campaign.

6 Validity checks

6.1 Balance

We use baseline data to assess the sample balance across treatment status. Table 1 shows that on average, 48% of children were girls and the average age of students was 4.5 years old. Among parents, 96% were women with an average age of 31 years old. Slightly more than 50% of the sample of parents had completed high school and 60% were head of the household. The average number of children at home was 2. Households have on average 7.32 assets. Internet access at home was 79% and 20% of parents were beneficiaries of a social program. Only 19% of parents in the sample reported that someone in their household was working remotely. Use of the government's remote education program AeC was very high, with 93% of parents reporting accessing it. Regarding emotional stress, 49% of parents reported experiencing at least one symptom of frequent stress in the past week. This indicator is 29% for child stress. Parents also reported that on average they engaged in 4.54 learning activities during the past three days and their child used 3.15 play-materials at home. Finally, 16% of the parents reported using some violent discipline practice. All these characteristics are balanced across treatment assignment groups, as expected from the randomization procedure.

6.2 Sample attrition

With respect to the baseline sample, attrition rates during the collection of the child learning assessment averaged 58.2%.⁸ These rates are similar to other telephone surveys conducted in developing countries during times of crisis (Henderson and Rosenbaum, 2020; Ballivian et al., 2015). The main reason for the loss of respondents was that parents changed their phone numbers

⁸Treated group: 58.6%, untreated group: 58.9% and pure control group: 57.6%

or the phone was deactivated. As reported on Table 2, our endline sample includes 1877 children for which there are available cognitive outcomes from the remote learning assessment (children sample). Table 3 shows that experimental groups in the children sample do not differ based on observable baseline characteristics.

We further explore sample attrition by estimating the probability that an observation is not in the children sample. Results are reported in Table 4. Reassuringly, attrition rates do not differ significantly across treatment arms, which suggests that attrition was not differential. Overall, some observable characteristics are correlated with attrition, although the size of the coefficients is very small in most cases. This sample selection does not affect the internal validity of the estimates but is relevant for extrapolating results. For example, the probability of not being in the children sample is positively correlated with whether the parent is head of the household, and whether someone at the household is working remotely.

6.3 Characteristics of the children's assessment

We collected data from parents to learn about their opinions in regards to their child's experience during the remote test. We also collected data from enumerators to understand their views about the process of applying the remote assessment, including their evaluation of the quality and difficulty of the application, whether any item did not work properly, the duration of the test and the number of times that parents interfered during the process. Table 5 shows that there are no statistically significant differences in these dimensions across groups, which implies that treatment effects are not explained by differences in the remote application of the test.

7 Results

Effects on children's cognitive skills. Examining the raw data first, Figure 4 plots the distributions of raw scores for the different treatment arms and their corresponding p-value from the Kolmogorov-Smirnov test of equality of distributions. Subfigures (a) - (c) compare the treated group to the control group. The distribution of numeracy scores of the treated group is clearly shifted to the right compared to the control group. Table 6 presents the results for our primary outcomes obtained from estimating equation (1) using the children sample. We find that children's

cognitive skills are 0.107 standard deviation (SD) higher in households assigned to receive the text message campaign compared to the control group. The coefficient is 0.116 SD when controlling for covariates. This result appears to be driven by numeracy skills (0.117 SD) rather than literacy skills. Table A.1 in the Appendix shows that effects are particularly strong for questions related to oral counting and number comparisons. There is some evidence of positive effects on questions related to sequences but the coefficient is barely significant.

These effects are sizable for a program that consisted on SMS only. Other low-cost intervention implemented during the first year of the pandemic that combined phone calls and text messages to primary school students in Botswana was associated with a 0.12 standard deviation increase in the ability to perform numerical operations (Angrist et al., 2020). Our effects are also substantial taking into account that the intervention lasted just 15 weeks. In fact, one possibility for the lack of effects on literacy skills might be that the duration of the intervention was too short. Evidence from the READY4K! text messaging program in the United States finds positive effects on early literacy skills (0.11 SD) but the program lasted eight months (York et al., 2019).

Overall, our results are consistent with research evidence that interventions providing timely and actionable information to parents can be effective in improving children's cognitive skills (Bergman, 2019). In a context where families' attention was limited by several issues related to the pandemic, the treated group is receiving a messaging campaign that is making children's education more salient, easier to remember and more simple to practice at home. In particular, since the control group received a shorter campaign compared to the treated group, the intervention is providing more frequent messages (biweekly vs. 3-4 weekly SMS), more messages in total (8 vs 54 SMS) and more relevant information (cognitive-specific information and activities vs existing information about AeC).

Spillover effects. Across all specifications in Table 6, we find no evidence of effects on children whose parents belonged to a treatment network but were not assigned to receive the text messages. This is consistent with subfigures (d) - (f) in Figure 4 that show no differences in the distributions of raw scores of the untreated group and the control group. Given the evidence of direct treatment effects, this absence of spillovers is somewhat surprising as priori we expected a stronger role of the networks to compensate for the lack of preschool instruction. However, one possibility is that we are

measuring impacts within networks that may not coincide with the groups where parents interact more actively. Note that the networks we observe are those defined by the preschool class that have a teacher in common, but schools closures in Costa Rica occurred very close to the beginning of the academic year, which probably did not give enough time to parents to get to know each other. Note that in our setting, spillover effects cannot be explained by peer effects between children in the classroom, as all children receive education only at home.

Treatment heterogeneity. We explore differential impacts by some baseline characteristics of the networks that we used for stratification. As reported in Table 7, we do not find strong evidence of heterogeneous impacts by size of the network. However, in Table 8 we find that direct impacts are much larger in networks with low baseline parent-teacher interactions, suggesting that messages are more effective for children in less connected households for whom the access to this information would have been hard to obtain otherwise. The absence of spillover effects is robust in all of these different subsamples.

8 Discussion

In this section we use some additional household data that was collected through an online follow-up survey.⁹ Parents in this sample responded to both the remote child learning test on the phone and the online survey sent through text message (parents sample). As shown in Panel A and Panel B of Table A.2, attrition rates were more severe, which may raise some further concerns about sample selection. Indeed, there is some differential attrition across treatment arms, even though differences are very small in magnitude (about 2 percentage points) and only significant at the 10% significance level. Table A.3 reports the estimated impacts on children cognitive outcomes using this sample. Compared to the previous results, the size of the coefficients is larger and there is some evidence of spillover effects on numeracy outcomes. One possible explanation for these differences is treatment

⁹The text message with the link to the survey was sent before a surveyor called to schedule the child's assessment. During the call, the surveyor reminded the parent to fill out the form. The survey captured information on the main mechanisms through which we expected the intervention might affect children's skills. These included parents' behaviors related to the learning environment at home, the engagement with the text message campaign and the government distance learning program. We collected information on activities parents performed at home with their children and measures of stress of students and parents. In addition, we measured parent's perception of child's abilities and their own parental self-efficacy. See Appendix B.1 for more details on how these indicators and scales were constructed.

effect heterogeneity where impacts are stronger for the sample of children in households of more responsive parents. In fact, Table A.4 shows that there are statistically significant differences in the level of education of parents and the proportion that benefits from a social program. There is also a higher proportion of mothers and a lower proportion of caregivers that are head of households which potentially implies more time available to spend with their child. This is also suggested by higher reports of child use of play-materials at home. However, another possibility is that these effects are confounding some selection bias as there is some evidence of small differential attrition across treatment groups. With these caveats in mind, we now explore potential mechanisms behind the effects on cognitive outcomes.

8.1 Parental communications and perspectives

We use parents' survey responses to understand the delivery and sharing of the text message campaign. As shown on Panel A of Table A.5, parents in the treated group were 47.8 percentage points more likely to recall having received the text messages as well as 46.7 percentage points more likely to recall the content of the messages than parents in the pure control group. Parents that belonged to a treatment network but were not assigned to receive the text messages did not report recalling the text messages more than the pure control group. This is somewhat expected as the untreated group and the pure control group received the same placebo campaign with 8 short messages reminding parents about AeC. Parents in the treated group were also more likely to identify the Ministry of Education as the sender (51.4 percentage points) and less likely to report receiving the messages from a teacher (4.5 percentage points). We don't find evidence that parents shared the text messages with the teacher, but we do find that a fraction of parents reported sharing the text messages with other parents (7.5 percentage points). However, there are no statistically significant differences in parents reports that they received the messages from other parents for any of the treatment arms.

We examine parental interactions on Panel B of Table A.5 and find that on average there are very few. Only 20% of the parents in the control group report communicating with other parent from preschool during the last week and there are no statistically significant differences across groups. Parents in the treated group are more likely to report that their children talk to other children from preschool (8 percentage points) compared to the control group (25%), which is an

indirect way in which parents might be interacting with each other. Finally, parents in the treated or untreated groups are not more likely to contact the teacher than parents in the control group (64%).

Panel C of Table A.5 shows that parents in the treated group also reported greater interest in that the AeC program continues when the child is back to in person instruction (5 percentage points). Interestingly, we find a negative effect for the untreated group on the report of parents accessing AeC resources from the government distance education program. As parents associate the SMS as being part of the Ministry of Education's AeC activities, this result suggests that parents that did not receive the text messages but learn about the program within their networks may have felt that their access to distance learning resources was limited. We do not find any differences in parents reports of whether the child is satisfied with the AeC program.

8.2 Parental involvement with the child

The intervention targeted parental engagement by providing parents with information about activities that can be easily performed at home to stimulate children's numeracy and literacy skills. The purpose was to take advantage of existing household objects and daily routines at home to make them productive for learning vocabulary, counting and building notions of quantity and relationships between quantities. Panel A of Table A.6 presents the results of the parents' survey for the number of activities parents reported performing with their children in the previous three days.¹⁰ Parents in the treated group report higher engagement with learning activities at home and higher involvement with the specific activities proposed in the SMS. There is also a small increase in the use of play-materials at home in the treated group. We did not find evidence of changes in the parental engagement of parents that belonged to treatment networks but were not assigned to receive the text messages.

The program included messages on positive parenting, time management, and healthy habits intended to help parents create a conducive home environment for learning. Panel B of Table A.6 shows that the increased involvement in learning activities did not translate into a greater sense of parental self-efficacy. There is neither evidence of changes in the use of violent discipline practices across treatment groups nor of the reported level of parent's or child's stress.

¹⁰See Appendix B.1 for more information about these scales.

Besides, an increase in time spent with children doing learning activities might allow parents to access to new information on their children's skills. Using survey information that asked parents to evaluate their child's skills in several dimensions, we constructed a dummy variable that indicates whether the child's location above or below the mean according to the standardized score coincides with the location above or below the mean according to the parents perceptions. Panel C of Table A.6 shows some evidence that parents in the treated group perceive the level of skills that their child has with a higher score compared to the control group. However, they are not more likely than the control group to evaluate whether their children are in the upper or lower half of the skills distribution.

9 Conclusion

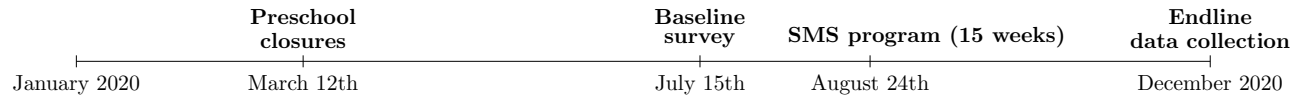
This paper shows evidence that a large-scale text message intervention can be effective in increasing preschool children's cognitive skills at home in a context where schooling is disrupted. We conducted a two-stage randomized experiment to quantify the direct and spillover short-term impacts of the program. The results show that after 15 weeks of intervention, cognitive skills increased by 0.11–0.12 standard deviations. The effect is explained mainly by an increase in numeracy skills. We do not find strong evidence that the text message campaign produced spillovers effects on untreated parents in treatment networks.

The intervention is low cost, with the cost per text message of about US\$0.014 (8 Colones). A total of 54 SMS were sent to 1,072 parents, leading to a total cost of US\$810. Given the average effect of 0.11 standard deviation increase in learning, the text message campaign is more cost-effective than other interventions aiming to increase early numeracy skills (Näslund-Hadley et al., 2018).

Our results are consistent with existing evidence that inexpensive behavioral interventions can empower parents to improve children's outcomes (Bergman, 2019) and further show that text message interventions offer a low-cost alternative for addressing school disruptions when children don't have access to in-person care. These findings have important implications for the provision of early childhood education in emergencies and conflict settings or in contexts where high-technology solutions are less scalable given the limited access to the Internet or digital devices.

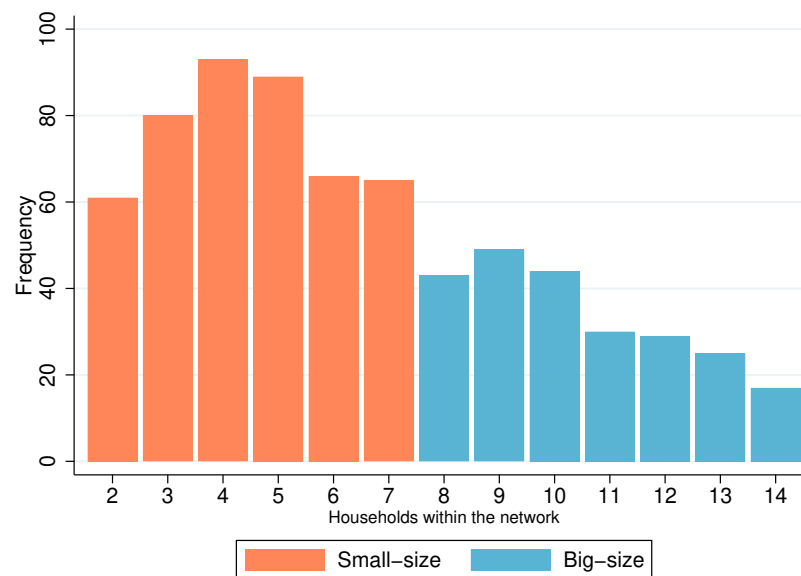
Figures

Figure 1: **Timeline**



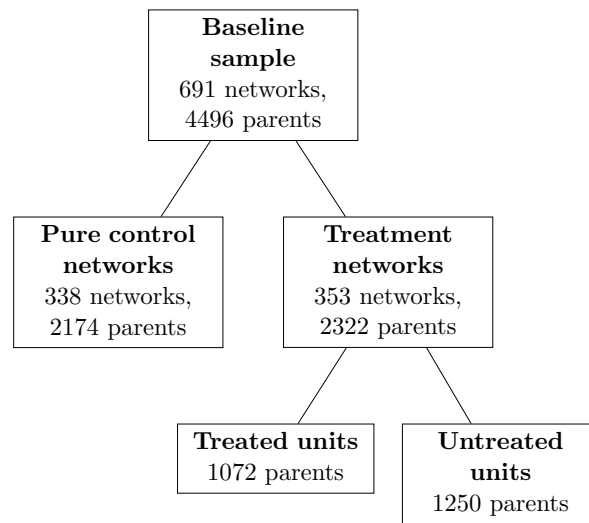
Notes: this figure provides a timeline of the project. The text message campaign was implemented over a period of 15 weeks (August 24–November 30, 2020). Baseline data was collected through an online household survey. Endline data was collected through a phone survey targeted to children.

Figure 2: **Distribution of network's size**



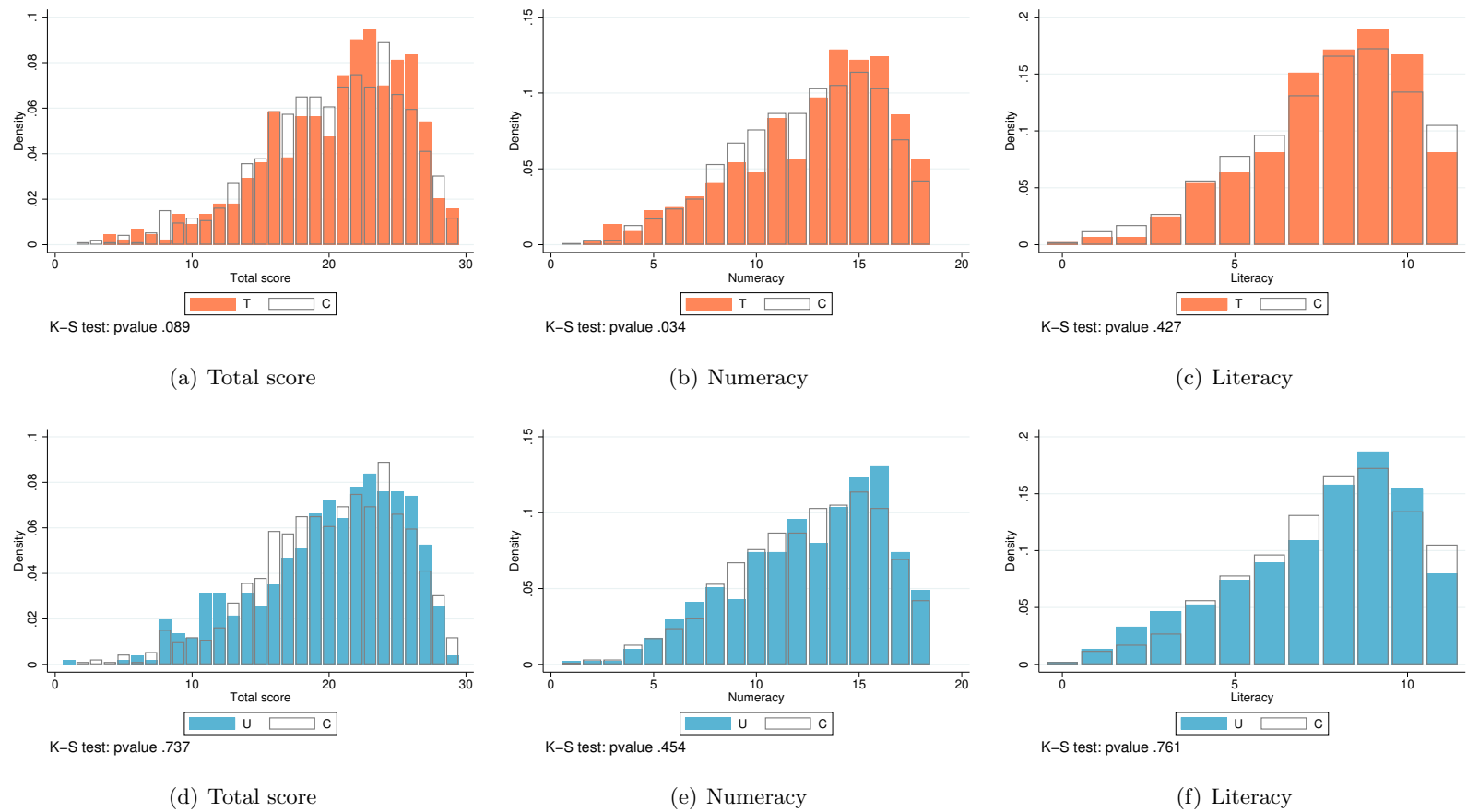
Notes: this figure shows the distribution of network's size in the experimental sample. Big-sized networks correspond to those with a number of parents higher than the median. We define a network as a group of parents that share the same preschool teacher.

Figure 3: **Experimental design**



Notes: this figure shows the two-stage experimental design. The randomization at the network level was stratified by province where the child center is located, network size, and the level of interaction between parents and teacher within the network. There are 28 strata. We define a network as a group of parents that share the same preschool teacher.

Figure 4: Distribution of cognitive skills (raw scores)



Notes: this figure plots the distributions of raw scores and the p-values of the corresponding Kolmogorov-Smirnov test. Subfigures (a) - (c) compare the treated group to the control group and subfigures (d) - (f) compare the untreated group to the control group.

Tables

Table 1: **Balance on baseline characteristics (experimental sample)**

	Assignment groups			Min.	Max.	P-value	Obs.
	Treated	Untreated	Pure control				
Child is female	0.47	0.48	0.48	0	1	0.860	4496
Child age	4.47	4.47	4.48	4	5	0.790	4496
Parent is female	0.96	0.96	0.96	0	1	0.512	4496
Parent age	31.00	31.13	31.17	18	71	0.788	4496
Parent completed high school	0.55	0.53	0.53	0	1	0.571	4496
Parent is head of household	0.61	0.61	0.60	0	1	0.731	4496
Children at home	2.02	1.98	1.99	1	8	0.437	4496
Household assets	7.32	7.32	7.36	0	13	0.818	4496
Internet access at home	0.78	0.78	0.79	0	1	0.545	4496
Parent is beneficiary of social program	0.19	0.21	0.21	0	1	0.537	4496
Remote work at home	0.18	0.18	0.19	0	1	0.659	4496
Household access AeC resources	0.93	0.92	0.93	0	1	0.484	4496
Parent stress	0.49	0.51	0.48	0	1	0.206	4496
Child stress	0.29	0.29	0.28	0	1	0.921	4496
Learning activities at home	4.54	4.53	4.53	0	5	0.928	4496
Play-materials at home	3.15	3.15	3.17	0	4	0.806	4496
Violent discipline	0.16	0.17	0.17	0	1	0.800	4496

Notes: this table reports summary statistics for baseline characteristics by treatment assignment group. The p-value corresponds to the F test that the coefficients of “Treated” and “Untreated” are globally non significant estimated from a linear regression using each baseline characteristic as the dependent variable. The index of household assets takes values between 0 and 13 and it is constructed as the summary score of 13 indicators. The variables for parent and child stress indicate whether parent reported frequent feelings of stress during the past week. The variable learning activities at home takes values between 0 and 5 and it is constructed as the summary score of five indicators adapted from UNICEF’s MICS6 Questionnaire for Children Under Five (UNICEF, 2020). The variable play-materials at home takes values between 0 and 4 and it is constructed as the summary score of 4 indicators adapted from UNICEF’s MICS6 Questionnaire for Children Under Five (UNICEF, 2020). The variable violent discipline indicates whether the parent has used any violent discipline practices according to UNICEF’s MICS6 Questionnaire for Children Under Five (UNICEF, 2020). All regressions include strata fixed effects. Standard errors are clustered at the network level. There are 1072 observations in the treated group, 1250 observations in the untreated group and 2174 observations in the pure control group.

Table 2: **Sample sizes**

	Baseline Household survey	Endline child assessment (Children sample)	Attrition rates
Total	4496	1877	58.2%
Treated	1072	443	58.6%
Untreated	1250	513	58.9%
Control	2174	921	57.6%

Notes: this table reports sample sizes for each survey. The children sample consists of observations for which there are available cognitive outcomes from the remote learning assessment.

Table 3: **Balance on baseline characteristics (children sample)**

	Assignment groups			Min.	Max.	P-value	Obs.
	Treated	Untreated	Pure control				
Child is female	0.49	0.50	0.51	0	1	0.811	1877
Child age	4.47	4.46	4.49	4	5	0.524	1877
Parent is female	0.97	0.94	0.97	0	1	0.127	1877
Parent age	30.83	31.46	31.24	18	71	0.272	1877
Parent completed high school	0.53	0.48	0.50	0	1	0.391	1877
Parent is head of household	0.59	0.62	0.57	0	1	0.114	1877
Children at home	2.10	2.09	2.03	1	7	0.463	1877
Household assets	7.31	7.20	7.25	1	13	0.702	1877
Internet access at home	0.80	0.76	0.80	0	1	0.124	1877
Parent is beneficiary of social program	0.22	0.27	0.23	0	1	0.133	1877
Remote work at home	0.17	0.15	0.15	0	1	0.711	1877
Household access AeC resources	0.93	0.93	0.94	0	1	0.234	1877
Parent stress	0.52	0.50	0.47	0	1	0.345	1877
Child stress	0.30	0.27	0.28	0	1	0.820	1877
Learning activities at home	4.51	4.49	4.49	0	5	0.867	1877
Play-materials at home	3.16	3.10	3.12	0	4	0.538	1877
Violent discipline	0.20	0.17	0.18	0	1	0.694	1877

Notes: this table reports summary statistics for baseline characteristics by treatment assignment group conditional on being in the children sample. The p-value corresponds to the F test that the coefficients of “Treated” and “Untreated” are globally non significant estimated from a linear regression using each baseline characteristic as the dependent variable. The index of household assets takes values between 0 and 13 and it is constructed as the summary score of 13 indicators. The variables for parent and child stress indicate whether parent reported frequent feelings of stress during the past week. The variable learning activities at home takes values between 0 and 5 and it is constructed as the summary score of five indicators adapted from UNICEF’s MICS6 Questionnaire for Children Under Five (UNICEF, 2020). The variable play-materials at home takes values between 0 and 4 and it is constructed as the summary score of 4 indicators adapted from UNICEF’s MICS6 Questionnaire for Children Under Five (UNICEF, 2020). The variable violent discipline indicates whether the parent has used any violent discipline practices according to UNICEF’s MICS6 Questionnaire for Children Under Five (UNICEF, 2020). All regressions include strata fixed effects. Standard errors are clustered at the network level. There are 443 observations in the treated group, 513 observations in the untreated group and 921 in the pure control group.

Table 4: **Attrition**

	(1) Attrited from children sample	(2) Attrited from children sample
Treated group	0.011 (0.019)	0.009 (0.018)
Untreated group	0.013 (0.019)	0.012 (0.019)
Child is female		-0.036** (0.015)
Child age		-0.004 (0.014)
Parent is female		-0.016 (0.036)
Parent age		-0.001 (0.001)
Parent completed high school		0.029* (0.016)
Parent is head of household		0.047*** (0.016)
Children at home		-0.019** (0.008)
Household assets		0.006 (0.005)
Internet access at home		-0.034* (0.019)
Parent is beneficiary of social program		-0.070*** (0.020)
Remote work at home		0.069*** (0.021)
Household access AeC resources		-0.041 (0.027)
Parent stress		0.000 (0.015)
Child stress		0.005 (0.018)
Learning activities at home		0.015* (0.008)
Play-materials at home		0.012 (0.009)
Violent discipline		-0.066*** (0.019)
Constant	0.576*** (0.012)	0.570*** (0.104)
Observations	4,496	4,496
Strata FE	28	28
Networks	691	691
F p-value	0.750	0.790
Mean dep. var	0.580	0.580

Notes: this table reports the coefficients of a model that estimates the probability that an observation attrited from the children sample. The model in column (2) controls for baseline characteristics including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. All regressions include strata fixed-effects. Standard errors clustered at the network level are reported in parenthesis. We define a network as a group of parents that share the same preschool teacher. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of "Treated group" and "Untreated group" are globally non significant.

Table 5: **Characteristics of the children’s assessment**

	Parents reports				Enumerators reports				
	(1) Child was comfortable	(2) Child enjoyed activity	(3) Questions were adequate	(4) Child understood questions	(5) Application difficulty	(6) Any item did not work	(7) Application quality	(8) Test duration (min)	(9) Total interruptions
Treated group	-0.014 (0.014)	0.006 (0.007)	0.025 (0.021)	0.026 (0.023)	-0.047 (0.053)	-0.011 (0.017)	0.039 (0.034)	-0.077 (0.252)	0.204 (0.471)
Untreated group	0.011 (0.011)	0.003 (0.007)	0.018 (0.020)	0.039* (0.021)	0.025 (0.048)	0.005 (0.018)	0.024 (0.031)	0.079 (0.235)	-0.457 (0.418)
Observations	1,877	1,877	1,877	1,877	1,877	1,877	1,877	1,877	1,877
Strata FE	28	28	28	28	28	28	28	28	28
Networks	640	640	640	640	640	640	640	640	640
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F p-value	0.220	0.680	0.440	0.160	0.460	0.710	0.490	0.870	0.410
Mean dep. var (control)	0.950	0.980	0.840	0.800	1.860	0.100	4.280	15.620	7.010
SD dep. var (control)	0.220	0.130	0.370	0.400	0.950	0.310	0.630	3.790	8.090

Notes: this table reports the estimated coefficients from model 1 using as dependent variables some characteristics of the remote child learning test. Variables in columns (1)-(4) were collected from parents and are measured as dummy variables as described in Appendix B.2. Variables in columns (5)-(9) were collected from enumerators. Column (5) is measured with a 5-point Likert scale (where 1 corresponds to very easy and 5 to very difficult). Column (6) is an indicator variable for whether any item from the test did not work properly during the application. Our measure for application quality in column (7) was computed as an average of the eight dimensions evaluated by enumerators as described in Appendix B.2. Column (8) corresponds to the test duration in minutes and column (9) to the number of times that parents interfered during the test. All regressions include strata fixed effects and control for baseline characteristics including: child’s gender; child’s age; parent’s gender; parent’s age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. Standard errors clustered at the network level are reported in parenthesis. We define a network as a group of parents that share the same preschool teacher. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of “Treated group” and “Untreated group” are globally non significant.

Table 6: **Effects on children's cognitive skills**

	(1)	(2)	(3)	(4)	(5)	(6)
	Composite score	Numeracy	Literacy	Composite score	Numeracy	Literacy
Treated group	0.107* (0.057)	0.108* (0.059)	0.071 (0.056)	0.116** (0.054)	0.117** (0.057)	0.080 (0.054)
Untreated group	0.010 (0.057)	0.065 (0.057)	-0.074 (0.057)	0.029 (0.055)	0.074 (0.056)	-0.046 (0.055)
Observations	1,877	1,877	1,877	1,877	1,877	1,877
Strata FE	28	28	28	28	28	28
Networks	640	640	640	640	640	640
Covariates	No	No	No	Yes	Yes	Yes
F p-value	0.150	0.160	0.060	0.090	0.100	0.100
Mean dep. var (control)	0.000	0.000	0.000	0.000	0.000	0.000
SD dep. var (control)	1.000	1.000	1.000	1.000	1.000	1.000

Notes: this table reports the estimated coefficients from model 1 using as dependent variables the standardized test scores collected from the remote child learning test. The model in columns (4) - (6) controls for baseline characteristics including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. All regressions include strata fixed effects. Standard errors clustered at the network level are reported in parenthesis. We define a network as a group of parents that share the same preschool teacher. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of "Treated group" and "Untreated group" are globally non significant.

Table 7: **Effect heterogeneity by size of the network**

	Big-sized networks			Small-sized networks		
	(1)	(2)	(3)	(4)	(5)	(6)
	Composite score	Numeracy	Literacy	Composite score	Numeracy	Literacy
Treated group	0.109 (0.072)	0.091 (0.077)	0.101 (0.072)	0.119 (0.083)	0.143* (0.086)	0.047 (0.082)
Untreated group	0.019 (0.073)	0.069 (0.077)	-0.061 (0.069)	0.040 (0.086)	0.080 (0.083)	-0.030 (0.090)
Observations	1,017	1,017	1,017	860	860	860
Strata FE	14	14	14	14	14	14
Networks	236	236	236	404	404	404
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
F p-value	0.310	0.420	0.110	0.360	0.240	0.680
Mean dep. var (control)	-0.020	-0.020	-0.030	0.030	0.020	0.030
SD dep. var (control)	0.990	0.990	1.000	1.010	1.010	1.000

Notes: this table reports the estimated coefficients from model 1 using as dependent variables the standardized test scores collected from the remote child learning test, separately by size of the network. We define a network as a group of parents that share the same preschool teacher. All regressions include strata fixed effects and control for baseline characteristics including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. Standard errors clustered at the network level are reported in parenthesis. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of "Treated group" and "Untreated group" are globally non significant.

Table 8: **Effect heterogeneity by level of parent-teacher interaction**

	High interaction			Low interaction		
	(1) Composite score	(2) Numeracy	(3) Literacy	(4) Composite score	(5) Numeracy	(6) Literacy
Treated group	0.076 (0.061)	0.074 (0.066)	0.055 (0.060)	0.266** (0.119)	0.280** (0.113)	0.164 (0.124)
Untreated group	0.040 (0.064)	0.094 (0.066)	-0.052 (0.064)	-0.015 (0.109)	0.002 (0.107)	-0.035 (0.115)
Observations	1,449	1,449	1,449	428	428	428
Strata FE	14	14	14	14	14	14
Networks	495	495	495	145	145	145
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
F p-value	0.460	0.290	0.290	0.030	0.020	0.250
Mean dep. var (control)	0.020	0.010	0.020	-0.060	-0.040	-0.070
SD dep. var (control)	0.970	0.980	0.980	1.100	1.080	1.060

Notes: this table reports the estimated coefficients from model 1 using as dependent variables the standardized test scores collected from the remote child learning test, separately for the level of parent-teacher interaction within the network. We define a network as a group of parents that share the same preschool teacher. All regressions include strata fixed effects and control for baseline characteristics including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. Standard errors clustered at the network level are reported in parenthesis. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of "Treated group" and "Untreated group" are globally non significant.

A Appendix: Tables and Figures

Table A.1: Effects on children's cognitive skills (specific tasks)

	(1) Spatial reasoning	(2) Oral counting	(3) Number comparisons	(4) Addition and subtraction	(5) Sequences	(6) Comparisons (weight and size)	(7) Syllabification	(8) Oral comprehension	(9) Expressive vocabulary
Treated group	0.081 (0.054)	0.124** (0.052)	0.118** (0.055)	0.031 (0.060)	0.095* (0.056)	0.004 (0.058)	0.027 (0.055)	0.064 (0.055)	0.096 (0.059)
Untreated group	0.064 (0.054)	0.016 (0.056)	-0.001 (0.056)	0.074 (0.056)	0.062 (0.056)	0.091* (0.054)	-0.076 (0.057)	0.049 (0.054)	-0.074 (0.055)
Observations	1,877	1,877	1,877	1,877	1,877	1,877	1,877	1,877	1,877
Strata FE	28	28	28	28	28	28	28	28	28
Networks	640	640	640	640	640	640	640	640	640
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F p-value	0.280	0.050	0.080	0.420	0.200	0.220	0.250	0.440	0.030
Mean dep. var (control)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SD dep. var (control)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Notes: this table reports the estimated coefficients from model 1 using as dependent variables the standardized test score collected from the remote child learning test, separately for each task. All regressions include strata fixed effects and control for baseline characteristics including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. Standard errors clustered at the network level are reported in parenthesis. We define a network as a group of parents that share the same preschool teacher. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of "Treated group" and "Untreated group" are globally non significant.

Table A.2: Validity checks for the parents sample

Panel A: Sample sizes and attrition rates

	Baseline Parents survey	Endline Child Assessment (Children sample)	Endline Parents survey	Endline Parents Survey (Parents sample)	Attrition rates
Total	4496	1877	1394	1090	75.7%
Treated	1072	443	302	249	76.7%
Untreated	1250	513	383	287	77.0%
Control	2174	921	709	554	74.5%

	(1) Attrited from parents sample	(2) Attrited from parents sample
Treated group	0.022 (0.016)	0.021 (0.016)
Untreated group	0.026* (0.015)	0.024* (0.014)
Observations	4,496	4,496
Strata FE	28	28
Networks	691	691
Covariates	No	Yes
F p-value	0.150	0.170
Mean dep. var	0.760	0.760

Panel B: Balance on baseline characteristics (parents sample)

	Assignment groups			Min.	Max.	P-value	Obs.
	Treated	Untreated	Pure control				
Child is female	0.50	0.51	0.53	0	1	0.900	1090
Child age	4.48	4.44	4.49	4	5	0.311	1090
Parent is female	0.95	0.96	0.98	0	1	0.231	1090
Parent age	31.31	31.43	31.45	18	59	0.852	1090
Parent completed high school	0.55	0.49	0.52	0	1	0.514	1090
Parent is head of household	0.53	0.59	0.56	0	1	0.320	1090
Children at home	2.00	2.14	2.05	1	7	0.202	1090
Household assets	7.42	7.20	7.26	2	12	0.468	1090
Internet access at home	0.79	0.77	0.81	0	1	0.572	1090
Parent is beneficiary of social program	0.22	0.28	0.25	0	1	0.449	1090
Remote work at home	0.19	0.15	0.14	0	1	0.283	1090
Household access AeC resources	0.92	0.92	0.96	0	1	0.007	1090
Parent stress	0.52	0.51	0.47	0	1	0.595	1090
Child stress	0.33	0.28	0.29	0	1	0.681	1090
Learning activities at home	4.55	4.52	4.52	0	5	0.795	1090
Play-materials at home	3.24	3.12	3.18	0	4	0.344	1090
Violent discipline practices	0.19	0.14	0.20	0	1	0.076	1090

Notes: This table reports validity checks for the sample with available information from the parents online endline survey (parents sample). Panel A reports sample sizes for each survey and the coefficients of a model that estimates the probability that an observation attrited from the parents sample. The model in column (2) controls for baseline characteristics including: child’s gender; child’s age; parent’s gender; parent’s age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. The p-value corresponds to the F test that the coefficients of “Treated group” and “Untreated group” are globally non significant. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. Panel B reports summary statistics for baseline characteristics by treatment assignment group conditional on being in the parents sample. The p-value corresponds to the F test that the coefficients of “Treated” and “Untreated” are globally non significant estimated from a linear regression using each baseline characteristic as the dependent variable. All regressions include strata fixed effects. Standard errors clustered at the network level are reported in parenthesis.

Table A.3: **Effects on children’s cognitive skills (parents sample)**

	(1)	(2)	(3)	(4)	(5)	(6)
	Composite score	Numeracy	Literacy	Composite score	Numeracy	Literacy
Treated group	0.196** (0.078)	0.206*** (0.078)	0.122 (0.076)	0.194** (0.076)	0.209*** (0.075)	0.113 (0.074)
Untreated group	0.070 (0.075)	0.140* (0.074)	-0.053 (0.076)	0.095 (0.073)	0.159** (0.072)	-0.028 (0.074)
Observations	1,090	1,090	1,090	1,090	1,090	1,090
Strata FE	28	28	28	28	28	28
Networks	547	547	547	547	547	547
Covariates	No	No	No	Yes	Yes	Yes
F p-value	0.040	0.020	0.090	0.030	0.010	0.170
Mean dep. var (control)	-0.010	-0.030	0.030	-0.010	-0.030	0.030
SD dep. var (control)	1.000	0.990	1.000	1.000	0.990	1.000

Notes: this table reports the estimated coefficients from model 1 using as dependent variables the standardized test scores collected from the remote child learning test. The sample includes children whose parents completed the online endline survey (parents sample). The model in columns (4) - (6) controls for baseline characteristics including: child’s gender; child’s age; parent’s gender; parent’s age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. All regressions include strata fixed effects. Standard errors clustered at the network level are reported in parenthesis. We define a network as a group of parents that share the same preschool teacher. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of “Treated group” and “Untreated group” are globally non significant.

Table A.4: **Baseline characteristics of the children sample separately by availability of parents survey**

	Is in parents sample		Diff. p-value	Obs.
	1	0		
Child is female	0.52	0.48	0.121	1877
Child age	4.48	4.48	0.684	1877
Parent is female	0.97	0.95	0.042	1877
Parent age	31.41	30.91	0.160	1877
Parent completed high school	0.52	0.48	0.033	1877
Parent is head of household	0.56	0.63	0.003	1877
Children at home	2.06	2.07	0.726	1877
Household assets	7.28	7.21	0.284	1877
Internet access at home	0.79	0.79	0.760	1877
Parent is beneficiary of social program	0.25	0.22	0.054	1877
Remote work at home	0.15	0.16	0.817	1877
Household access AeC resources	0.94	0.93	0.239	1877
Parent stress	0.49	0.49	0.852	1877
Child stress	0.30	0.27	0.252	1877
Learning activities at home	4.53	4.45	0.155	1877
Play-materials at home	3.18	3.04	0.001	1877
Violent discipline practices	0.18	0.19	0.748	1877

Notes: this table reports summary statistics for baseline characteristics in the children sample separately by whether observations have available information from the parents sample. The p-value corresponds to the test that the difference between the groups is equal to zero estimated from a linear regression using each baseline characteristic as the dependent variable. The index of household assets takes values between 0 and 13 and it is constructed as the summary score of 13 indicators. The variables for parent and child stress indicate whether parent reported frequent feelings of stress during the past week. The variable learning activities at home takes values between 0 and 5 and it is constructed as the summary score of five indicators adapted from UNICEF's MICS6 Questionnaire for Children Under Five (UNICEF, 2020). The variable play-materials at home takes values between 0 and 4 and it is constructed as the summary score of 4 indicators adapted from UNICEF's MICS6 Questionnaire for Children Under Five (UNICEF, 2020). The variable violent discipline indicates whether the parent has used any violent discipline practices according to UNICEF's MICS6 Questionnaire for Children Under Five (UNICEF, 2020). All regressions include strata fixed effects. Standard errors are clustered at the network level.

Table A.5: Parental communications and perspectives

Panel A: Delivery and sharing of SMS							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Received SMS	Remembered SMS content	Received SMS from MEP	Received SMS from teacher	Received SMS from other parent	Shared SMS with other parent	Shared SMS with teacher
Treated group	0.478*** (0.031)	0.467*** (0.036)	0.514*** (0.031)	-0.045** (0.021)	-0.006 (0.004)	0.075*** (0.022)	0.008 (0.016)
Untreated group	-0.028 (0.035)	-0.018 (0.027)	-0.025 (0.033)	-0.014 (0.022)	0.005 (0.007)	-0.013 (0.016)	-0.010 (0.014)
Observations	1,090	1,090	1,090	1,090	1,090	1,034	989
Strata FE	28	28	28	28	28	28	28
Networks	547	547	547	547	547	531	518
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F p-value	0.000	0.000	0.000	0.090	0.090	0.000	0.540
Mean dep. var (control)	0.410	0.170	0.360	0.110	0.010	0.050	0.030
SD dep. var (control)	0.490	0.370	0.480	0.310	0.070	0.210	0.180

Panel B: Interactions			
	(1)	(2)	(3)
	Parent talked to other parent from preschool	Child talked to other child from preschool	Parent contacted the teacher
Treated group	0.000 (0.031)	0.080** (0.035)	0.028 (0.037)
Untreated group	-0.027 (0.028)	0.030 (0.033)	-0.016 (0.034)
Observations	1,090	1,090	1,090
Strata FE	28	28	28
Networks	547	547	547
Covariates	Yes	Yes	Yes
F p-value	0.570	0.070	0.540
Mean dep. var (control)	0.190	0.250	0.640
SD dep. var (control)	0.390	0.430	0.480

Panel C: National remote learning program *Aprendo en Casa* (AeC)

	(1)	(2)	(3)
	Parent wants AeC to continue	Child accessed AeC	Child satisfied with AeC
Treated group	0.050** (0.020)	0.009 (0.024)	0.002 (0.030)
Untreated group	0.033 (0.021)	-0.076*** (0.029)	0.006 (0.030)
Observations	1,090	1,083	939
Strata FE	28	28	28
Networks	547	546	500
Covariates	Yes	Yes	Yes
F p-value	0.030	0.010	0.980
Mean dep. var (control)	0.880	0.890	0.830
SD dep. var (control)	0.320	0.320	0.380

Notes: This table reports the estimated coefficients from model 1 using as dependent variables several dichotomous variables of parental communications and perspectives. In Panel A: column (1) indicates whether parent reports receiving SMS from the program; column (2) indicates whether parent recalls the content from the SMS; columns (3)-(5) indicates whether parent reports receiving the SMS from the MEP, teacher or a parent, respectively and columns (6)-(7) indicates whether parent reports sharing the SMS with other parent or the teacher, respectively. In Panel B: column (1) indicates whether the parent reports talking to another parent from the preschool over the last week; column (2) indicates whether the parent reports that the child talked to another child from the preschool and column (3) indicates whether the parent reports contacting the teacher in the last week. In Panel C: column (1) indicates whether the parent considers helpful that the resources from AeC continue when children get back to preschool; column (2) indicates whether the parent reported that their child accessed the resources from AeC and column (3) indicates whether the parent reported that their child is very satisfied or satisfied with the resources from AeC. All regressions include strata fixed effects and control for baseline characteristics including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. Standard errors clustered at the network level are reported in parenthesis. We define a network as a group of parents that share the same preschool teacher. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of "Treated group" and "Untreated group" are globally non significant.

Table A.6: Parental involvement with the child

Panel A: Engagement with learning practices				
	(1)	(2)	(3)	
	Learning activities at home	SMS proposed activities	Play-materials at home	
Treated group	0.144** (0.057)	0.388*** (0.117)	0.091* (0.052)	
Untreated group	0.024 (0.055)	0.155 (0.117)	0.023 (0.053)	
Observations	1,090	1,090	1,090	
Strata FE	28	28	28	
Networks	547	547	547	
Covariates	Yes	Yes	Yes	
F p-value	0.030	0.000	0.220	
Mean dep. var (control)	4.560	4.140	3.320	
SD dep. var (control)	0.880	1.680	0.840	

Panel B: Socio-emotional skills				
	(1)	(2)	(3)	(4)
	Parent self-efficacy	Violent discipline	Parent stress	Child stress
Treated group	0.002 (0.024)	-0.001 (0.023)	0.022 (0.034)	0.006 (0.022)
Untreated group	-0.001 (0.022)	0.029 (0.025)	0.042 (0.033)	-0.013 (0.019)
Observations	1,090	1,090	1,090	1,090
Strata FE	28	28	28	28
Networks	547	547	547	547
Covariates	Yes	Yes	Yes	Yes
F p-value	1.000	0.480	0.430	0.700
Mean dep. var (control)	3.480	0.160	0.320	0.090
SD dep. var (control)	0.310	0.360	0.470	0.290

Panel C: Parent's beliefs		
	(1)	(2)
	Parent's perception of child's skills	Child's relative position and parent's perception coincide
Treated group	0.049** (0.025)	0.053 (0.037)
Untreated group	-0.001 (0.024)	0.051 (0.033)
Observations	1,090	1,090
Strata FE	28	28
Networks	547	547
Covariates	Yes	Yes
F p-value	0.120	0.180
Mean dep. var (control)	2.370	0.650
SD dep. var (control)	0.320	0.480

Notes: This table reports the estimated coefficients from model 1 using as dependent variables several measures of parental involvement with the child. Panel A focus on engagement with learning practices: learning activities at home, SMS proposed activities and use of play materials. Panel B focus on socio-emotional skills: parent self-efficacy, violent discipline, parent stress and child stress. Panel C focus on parent's beliefs about the child's performance and an indicator variable of whether the child's location above or below the mean according to the standardized score coincides with the location above or below the mean according to the parents perceptions. All regressions include strata fixed effects and control for baseline characteristics including: child's gender; child's age; parent's gender; parent's age; whether the parent completed high-school; whether parent is head of household; number of children at home; number of household assets; Internet access; whether the parent is beneficiary of a social program; whether there is remote work at home; household access to AeC; parent stress; child stress; learning activities practiced at home; use of play-materials at home and use of violent discipline. Standard errors clustered at the network level are reported in parenthesis. We define a network as a group of parents that share the same preschool teacher. Significance at the one, five and ten percent levels is indicated by ***, ** and *, respectively. The p-value corresponds to the F test that the coefficients of "Treated group" and "Untreated group" are globally non significant.

B Appendix: Data sources and variables' construction

B.1 Main variables collected through online surveys

- **Parent stress:** this variable indicates whether parent reported feeling stressed frequently during the past week. We used a modified version of the Center for Epidemiologic Studies Depression Scale Revised(CESD-R, 2020). This dummy equals to 1 if parent reported either 3 or 4 in any item. This variable is available for baseline and endline.

Below is a list of the ways you might have felt or behaved.

Please check the boxes to tell me how often you have felt this way in the past week or so:

	Never/rarely (less than a day) [value = 1]	Sometimes (1-2 days) [value = 2]	Often (3-4 days) [value = 3]	Most of the time (5-7 days) [value = 4]
1. Feel tired or without energy				
2. Have trouble falling sleep				
3. Not have appetite				
4. Feel sad/ depressed				
5. Feel nervous/ worried				

- **Child stress (reported by parent):** this variable indicates whether parent reported that their child was feeling stressed frequently during the past week. We used a modified version of the Center for Epidemiologic Studies Depression Scale Revised(CESD-R, 2020). This dummy equals to 1 if parent reported either 3 or 4 in any item. This variable is available for baseline and endline.

In the past week or so, did you notice that your child?

	Never/rarely (less than a day) [value = 1]	Sometimes (1-2 days) [value = 2]	Often (3-4 days) [value = 3]	Most of the time (5-7 days) [value = 4]
1. Was nervous/ tense				
2. Worried too much				
3. Was sad				
4. Could not sleep well				

- **Play-materials at home:** this variable takes values between 0 and 4. Items were adapted from UNICEF's MICS6 Questionnaire for Children Under Five(UNICEF, 2020). This index is constructed as the summary score of the following four indicators. This variable is available for baseline and endline.

Does your child play with?

	Yes [value = 1]	No [value = 0]
1. Household objects or objects found outside (pots, rocks, sticks)		
2. Store-bought toys		
3. Homemade toys		
4. Technology (smartphone, tablets, computer)		

- **Learning activities at home:** this variable takes values between 0 and 5. Items were adapted from UNICEF's MICS6 Questionnaire for Children Under Five(UNICEF, 2020). This index is constructed as the summary score of the following five indicators. This variable is available for baseline and endline.

In the past three days, did you or any other household member (older than 15 years old) engage in any of the following activities with the child?

	Yes [value = 1]	No [value = 0]
1. Read books /look at pictures		
2. Tell stories		
3. Sing songs		
4. Play		
5. Name objects or draw things		

- **SMS proposed activities:** this variable takes values between 0 and 6. This index is constructed as the summary score of the following six indicators. This variable is available for endline.

In the past three days, did you or any other household member (older than 15 years old) engage in any of the following activities with the child?

	Yes [value = 1]	No [value = 0]
1. Do exercise		
2. Count objects		
3. Compare objects "bigger/smaller than"		
4. Play using adding and subtraction games		
5. Name objects from the household/food/animals		
6. Play separating words into syllables		

- **Violent discipline:** this dummy variable indicates whether parent has used any of the violent discipline practices (3, 4, 6, 7, 8, 9 or 10). Items were adapted from UNICEF's MICS6 Questionnaire for Children Under Five(UNICEF, 2020). This variable is available for baseline and endline.

Have you or other adult in your household used this method with the child in the past month?

	Yes [value = 1]	No [value = 0]
1. Took away privileges, forbade something		
2. Explain why behavior was wrong		
3. Shook him/her		
4. Yelled at him/her		
5. Gave him/her something else to do		
6. Spanked him/her with bare hand		
7. Hit him/her with a belt or hard object		
8. Called him/her dumb or lazy		
9. Hit him/her on the face		
10. Hit him/her on the hand		

- **Household assets:** This index takes values between 0 and 13, it is constructed as the summary score of the following indicators. This variable is available for baseline.

Do you have any of the following services/goods at home?

	Yes [value = 1]	No [value = 0]
1. Bathroom		
2. Refrigerator		
3. Motorcycle		
4. Car		
5. Piped indoor water		
6. Electricity		
7. Air conditioner		
8. Colour TV		
9. Radio		
10. Telephone		
11. Smartphone		
12. Computer/laptop		
13. Tablet		

- **Parent's perception of child's skills:** This index takes values between 1 and 3, it is constructed as the average value of the following sixteen items. This variable is available for endline. Using this information we also constructed an indicator of whether the child's location above or below the mean according to the standardized score coincides with the location above or below the mean according to the parent perception.

How would you evaluate your child's skills in the following tasks?

	Low [value = 1]	Medium [value = 2]	High [value = 3]
1. Counting more than 10 objects			
2. Counting more than 30 objects			
3. Distinguishing between "something big" and "something small"			
4. Distinguishing between "more objects" and "less objects"			
5. Addition exercises			
6. Subtraction exercises			
7. Writing some letters or numbers			
8. Drawing triangles and other geometric figures			
9. Knowing the name of food/kitchen/bathroom items			
10. Pronouncing words clearly			
11. Telling stories using full sentences			
12. Using future tense correctly			
13. Saying his/her name and address			
14. Understanding stories or simple instructions			
15. Using plurals and past tense correctly			
16. Talking about his/her own daily life			

- **Parent self-efficacy:** This index takes values between 1 and 4, it is constructed as the average value of the following eleven items. Some items were taken from Gibaud-Wallston & Wandersman's Parenting Sense of Competence Scale (Johnston and Mash, 1989). This variable is available for endline.

How much do you agree with the following statements?

	Strongly disagree [value = 1]	Disagree [value = 2]	Agree [value = 3]	Strongly agree [value = 4]
1. I feel prepared to support my child's education				
2. Sometimes I react too strongly when my child misbehaves	[value = 4]	[value = 3]	[value = 2]	[value = 1]
3. I can help my child to learn new things				
4. I understand my child's feelings				
5. I can construct a happy and peaceful home				
6. I can control my child's emotions				
7. Being a parent is manageable, any problems are easily solved				
8. I can explain things to my child				
9. I can follow routines with my child				
10. I frequently tell my child I love him/her				
11. I congratulate my child for doing things right				

- **Teacher reported high interaction with parents:** This indicator takes value 1 if teacher reported either 4 or 5. This classifies networks as "high parent-teacher interaction" or "low parent-teacher interaction". This variable is available for baseline.

About the parents in your class:

	Nobody [value = 1]	Some [value = 2]	Half [value = 3]	Almost all [value = 4]	All [value = 5]
How many families have communicated with you during this time?					

B.2 Remote assessment of early childhood cognitive skills

We designed an instrument to measure cognitive ability in early childhood based on a remote adaptation of existing standardized tests such as the Early Grade Mathematics Assessment EGMA (RTI International, 2009) and the Measuring Learning Quality and Outcomes (UNESCO et al., 2017)¹¹. These tests provide relevant and valid content to capture key cognitive skills in the early childhood. We complemented this information with insights from officials and early childhood experts at MEP to align the instrument with specific requirements of the curriculum and the local context. Our telephone-based instrument is intended to be used as a low-stakes assessment to monitor early cognitive ability and not as a screening test for diagnostic use. In addition, we specifically tailor our instrument to account for the restriction that, due to poor internet connectivity in the areas of interest, the cognitive tasks could not rely on visual aids.¹²

Data collection tools. We used *SurveyCTO* to design and construct a survey form in ODK language that allows to collect data from a mobile app. The advantages of this tool include real time collection and monitoring, the possibility of adding audio and text audits for quality verification and plug-ins to integrate mobile services with the *SurveyCTO* app. We provided specific mobile devices and headphones to the field staff to ensure that they all had the recommended technology for the activities. The average test was 15.63 minutes long with a maximum of 58.33 minutes and a minimum of 4.41 minutes.

Scores. The scale contains 29 items grouped into 10 tasks: spatial reasoning, oral counting, number comparisons, addition, expressive vocabulary, syllabication, sequences, subtraction, weight and size comparisons and oral comprehension. We treat items as binary in the sense that each of them can be either correct (score = 1) or incorrect (score = 0) and compute a composite score as the sum of all the items giving each item the same weight. This implies that each item is equally important to each other and that all items are positively related to cognitive ability. Missing responses when the child did not know the answer are counted as incorrect. The average raw composite score was 20.20 (std. dev. = 5.16, min = 1, max = 29). We also compute a numeracy score and a literacy score as the sum of items that are related to each knowledge area. These are spatial reasoning, oral counting, number comparisons, addition, subtraction, sequences and weight and size comparisons for numeracy and syllabication, oral comprehension and expressive vocabulary for literacy. The average numeracy raw score was 12.57 (std. dev. = 3.54, min = 1, max = 18) and the average literacy raw score was 7.63 (std. dev. = 2.38, min = 0, max = 11). Each of these scores was standardized to have a mean of zero and a standard deviation of one in the pure control group.

Validity and reliability. The Cronbach's alpha coefficient across the 29-item scale is 0.8217, suggesting that items have high internal consistency to measure cognitive ability. We complement this evidence with Item Response Theory to characterize items in terms of their difficulty/location and discrimination/information parameters by fitting a two-parameter logistic (2PL) model¹³. Figure (a) shows that there are four items about number comparisons, sequences and listening comprehension that are relatively difficult (positive estimates). The items that are the least difficult are about weight and size comparisons, oral comprehension and spatial reasoning. In terms of discrimination, the items that can distinguish more between low and high ability children are about syllabication and sequences. Those with the lowest discrimination are about listening comprehension and number comparison. Figure (b) shows the expected score for different levels of the estimated children's ability: children with above-average ability are expected to have composite score above 21 and about 95% of randomly selected children are expected to score between 9.77 and 27.

Enumerators. The assessments were applied by trained enumerators with relevant experience with technology

¹¹ In 2020, we worked with the Ministry of Education of Peru and the *Instituto Colombiano de Bienestar Familiar* (ICBF) in Colombia on a pilot to test a remote version of the MELQO questionnaires as a tool to measure early childhood development. These questionnaires were a useful input for the final version of the instruments used in this study.

¹² For example, we don't include items related to writing, mental transformation, letters', numbers' or shapes' identification.

¹³ For estimation we use the Stata package for IRT.

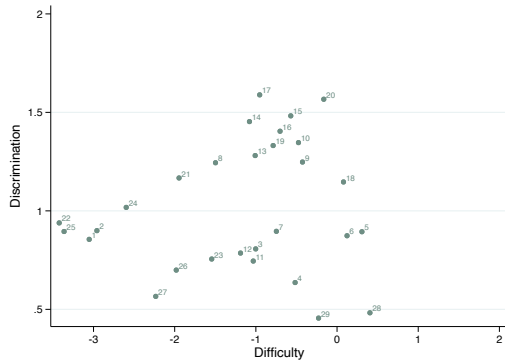


Figure (a): 2PL Item Parameters

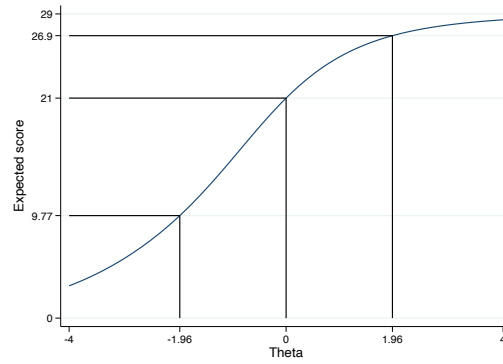


Figure (b): Test Characteristic Curve

and data collection, as well as previous experience working with children during interviews, test application or teaching. Enumerators shared their opinions about the difficulty of the application itself, 75.81% considering it easy or very easy to collect and 90% reporting that all items worked properly during the application. We also asked enumerators to evaluate the quality of the test application in several dimensions using a Likert scale from 1 to 5 (where 1 is very low and 5 is very high). Their feedback shows an overall satisfaction with the assessment's administration in terms of the call quality, the communication with the caregivers and the child and the environment of the home during the assessment (background noise, distractions from people in the room, child's attitude, etc.).

	Mean	Std. Dev.	Min	Max
Call quality	4.494	0.774	1	5
Communication with caregiver	4.746	0.505	1	5
Communication with child	4.368	0.814	1	5
Parent behavior	4.054	0.965	1	5
Child understood instructions	4.117	0.891	1	5
Child understood activities	4.103	0.890	1	5
Child attention	4.170	0.906	1	5
Environment at home	4.226	0.935	1	5
Observations	1877			

Caregivers. We asked caregivers to be near the child and monitor the phone call while the call was taken place. This implies a very important role for them during the administration of the test. Enumerators required caregivers to find a place without noise or distractions for the call, make sure they are in a place with a stable signal for the call and verify that the device has enough power charge for the call. Then, caregivers were told to put the cell phone on speaker and make sure the child is seated in front of the device and hands free, without touching the device. Importantly, caregivers should have stand behind the child and avoid interrupting or helping answering the questions. To mitigate potential caregiver intervention, the enumerator encouraged the caregiver to allow the child to answer each of the questions by himself/herself and clarified that caregivers are not supposed to interrupt to provide answers, repeat the questions or encourage their child. In addition, we included reminder messages during the call when caregivers were not behaving appropriately. Whenever there was any interruption in the first three items of the test the enumerator said to the caregivers “remember not to provide answers to your child or repeat the questions or encourage him to answer. The most important thing here is his own effort”. Other reminders said “The best way to help your child is letting him to respond by himself”, “Do not worry because this activity does not have any grading” and “Your child is doing a very good job”. Despite these strategies, we found that 81.35% of caregivers interrupted at least once during the test and the average number of interruptions was 8.5 times.

We gathered qualitative information from caregivers and overall, results were very positive in terms of the child being comfortable during the assessment and the child enjoying the activity. In addition, caregivers considered that the questions asked to the child were adequate for their age-group and believe that their child understood the questions.

	Mean	Std. Dev.	Min	Max
Child was comfortable	0.946	0.226	0	1
Child enjoyed activity	0.986	0.119	0	1
Questions were adequate	0.853	0.354	0	1
Child understood the questions	0.815	0.388	0	1
Observations	1877			

References

- Angrist, N., Bergman, P., and Matsheng, M. (2020). School's out: Experimental evidence on limiting learning loss using "low-tech" in a pandemic. *National Bureau of Economic Research*.
- Baird, S., Bohren, J. A., McIntosh, C., and Özler, B. (2018). Optimal design of experiments in the presence of interference. *Review of Economics and Statistics*, 100(5):844–660.
- Ballivian, A., Azevedo, J. P., and Durbin, W. (2015). *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*, chapter Using Mobile Phones for High-Frequency Data Collection, pages 21–39. London: Ubiquity Press.
- Barrera, O., Macours, K., Premand, P., and Vakis, R. (2020). Texting parents about early child development : Behavioral changes and unintended social effects. *Policy Research Working Paper; No. 9492*. World Bank, Washington, DC.
- Bergman, P. (2019). How behavioral science can empower parents to improve children's educational outcomes. *Behavioral Science & Policy*, 5(1):52–67.
- Bouchane, K. et al. (2019). Early childhood development and early learning for children in crisis and conflict. *UNESCO Global Education Monitoring Report*.
- CESD-R (2020). The center for epidemiologic studies depression scale (cesd). <https://cesd-r.com/about-cesdr/>.
- Cortes, K. E., Fricke, H., Loeb, S., Song, D. S., and York, B. N. (2021). Too little or too much? actionable advice in an early-childhood text messaging experiment. *Education Finance and Policy*, 16(2):209–232.
- Doss, C., Fahle, E. M., Loeb, S., and York, B. N. (2019). More than just a nudge supporting kindergarten parents with differentiated and personalized text messages. *Journal of Human Resources*, 54(3):567–603.
- Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 118(3):815–842.
- Elango, S., García, J. L., Heckman, J. J., and Hojman, A. (2015). Early childhood education. In *Economics of Means-Tested Transfer Programs in the United States, Volume 2*. University of Chicago Press.

- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*.
- Henderson, S. and Rosenbaum, M. (2020). Remote surveying in a pandemic: research synthesis. *Innovation for Poverty Action*.
- Hurwitz, L. B., Lauricella, A. R., Hanson, A., Raden, A., and 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.
- Johnston, C. and Mash, E. J. (1989). A measure of parenting satisfaction and efficacy. *Journal of Clinical Child Psychology*, 18(2):167–175.
- Mayer, S. E., Kalil, A., Oreopoulos, P., and Gallegos, S. (2019). Using behavioral insights to increase parental engagement the parents and children together intervention. *Journal of Human Resources*, 54(4):900–925.
- Moffit, R. (2001). Policy interventions, low-level equilibria and social interactions. In Durlauf, S. N. and Young, P., editors, *Social Dynamics*, pages 45–82. MIT Press.
- Näslund-Hadley, E., Bando, R., and Gertler, P. (2018). Inquiry and problem based pedagogy: Evidence from 10 field experiments. Technical report, Inter-American Development Bank.
- Näslund-Hadley, E., Hernández Agramonte, J. M., Montaña, K., Namen, O., Alpizar, G., Luna, Úrsula; Ochoa, L., García, J. F., Peña de Osorio, B., Biehl, M. L., Maragall, J., Méndez, C., and Thompson, J. (2020). Remote initial education and mental health during the covid-19 pandemic. Technical report, Education Policy Brief 4, Latin American and the Caribbean, Education Division, Social Sector. Washington, DC: Inter-American Development Bank.
- OECD (2017). *Education in Costa Rica, Reviews of National Policies for Education*. OECD Publishing, Paris.
- Ponguta, L. A., Moore, K., Varghese, D., Hein, S., Ng, A., Alzaghoul, A. F., Benavides Camacho, M. A., Sethi, K., and Al-Soleiti, M. (2022). Landscape analysis of early childhood development in emergencies. *Inter-agency Network for Education in Emergencies*.
- RTI International (2009). Early Grade Mathematics Assessment (EGMA): A Conceptual Framework Based on Mathematics Skills Development in Children.
- UNESCO, UNICEF, Institution, B., and Bank, W. (2017). Overview: MELQO: Measuring Early Learning Quality and Outcomes.

- UNICEF (2020). Unicef multiple indicator cluster surveys (mics). <https://mics.unicef.org/tools>.
- Vazquez-Bare, G. (2022). Identification and estimation of spillover effects in randomized experiments. *Journal of Econometrics*.
- York, B. N., Loeb, S., and Doss, C. (2019). One step at a time the effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, 54(3):537–566.