Policy Research Working Paper 10413

Improving Enrollment and Learning through Videos and Mobiles

Experimental Evidence from Northern Nigeria

Victor Orozco-Olvera
Ericka Rascon-Ramirez

WORLD BANK GROUP
Development Economics
Development Impact Evaluation Group
April 2023

Policy Research Working Paper 10413

Abstract

In northern Nigeria, half of primary school-age children attend school, half of girls are married before turning 15, and one in five people can read a whole sentence. Conducted in rural, low literate communities governed by traditional norms, this paper presents the results of a cluster randomized controlled trial that tested community screenings to reshape parental aspirations and attitudes toward education, and as a reinforcing arm, the distribution of mobiles with engaging apps to teach 6-9-year-old children to read. Twelve months after the screenings, children were 42 percent less likely to be out of school, but as expected, their learning levels did not improve. In the communities that were provided the mobile reinforcer, literacy and numeracy skills increased by 0.46 and 0.63 standard deviation, respectively. The impacts of the combined intervention on school attendance and learning gains were similar for boys and girls. For non-targeted older siblings, the intervention increased learning by 0.34 and 0.47 standard deviation and reduced the likelihood of teenage pregnancy and early entrance into the labor market by 13 and 14 percent, respectively. The mechanisms behind these effects include improved parental aspirations and expectations, improved attitudes and social norms, higher self-efficacy beliefs of parents, and increased time for home learning activities. Relative to other educational investments that have been evaluated in developing countries, the combined intervention is highly effective and cost-effective.

This paper is a product of the Development Impact Evaluation Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at vorozco@worldbank.org and e.ramirez@mdx.ac.uk.

Improving Enrollment and Learning through Videos and Mobiles: Experimental Evidence from Northern Nigeria*

Victor Orozco-Olvera, *DIME, The World Bank*
Ericka Rascón-Ramírez, *Middlesex University*

Keywords: school enrollment, aspirations, learning, disruptive technologies, self-efficacy, early marriage, early parenthood, gender norms. **JEL Classification:** I21, J12, J13, J16, L82, O15, Z13.

*This research is part of the entertainment-education program of the World Bank's Development Impact Evaluation Department (DIME). Computational replicability was verified by DIME Analytics. We thank the Nigeria Federal and the State ministries of education of Kano and Jigawa, in particular Dr. Folake Olatunji-David, for supporting the testing of innovations and for facilitating access to local communities. We thank Impact(Ed) International, Curious Learning, Global Digital Library, Breaking Barriers in Development, and MTN for producing the interventions' materials or supporting their delivery. We are immensely grateful to our field coordinator Edwin Ikhuoria and Hanovia Ltd, the survey firm that conducted the data collection. Laura Costica and Tomas Koutecky provided excellent research assistance. We thank seminar participants at the Royal Economic Society, Centro de Investigación y Docencia Económicas (CIDE), Universidad de Granada, Universidad de Sevilla, University of Maryland, and World Bank DIME. This project was funded by the World Bank Nigeria Partnership for Education Project, the World Bank Research and Support Budget, and the Norwegian Agency for Development Cooperation. This study received ethical clearance from Solutions IRB (IORG0007116) and was preregistered at the AEA RCT Registry `https://www.socialscienceregistry.org/trials/3619` prior to baseline data collection. The views expressed in this article are those of the authors and do not necessarily reflect the views of their institutions. All errors are the authors' sole responsibility.

1 Introduction

Despite important progress in school attendance in developing countries, there are 61 million children out-of-school in the world (UNESCO 2019). Many children who attend school often complete primary school without knowing how to read and write (Banerjee and Duflo 2011). This learning crisis is widespread in developing countries. Prior to the COVID-19 pandemic, 86 percent of ten-year-old children in Sub-Saharan Africa were learning-poor, defined by the World Bank as not being able to read an age-appropriate text. Disruptions to school systems caused by the pandemic have further worsened learning poverty across developing countries (The World Bank 2022).

While part of the learning crisis is a result of school-based constraints such as the lack of adequate school infrastructure (Duflo 2004), considerable heterogeneity in learning levels of students within the same grade (Muralidharan et al. 2019), and teachers lacking the skills or motivation to teach effectively (Popova et al. 2018), demand-side constraints are also important drivers behind this crisis. In low-income settings, parents often have low aspirations for their children's educational achievement and overall futures, which in turn influence present schooling and early marriage decisions. Illiterate parents may also have low self-efficacy beliefs about their ability to support their children's learning. The impact of these constraints is worsened by a lack of learning materials in the child's mother tongue (Brunette et al. 2019). Consequently, demand-side constraints, chronically present in rural and low-literacy environments, mean that children start primary school unprepared or are not given the opportunity to even been enrolled. This, in turn, leads to limited learning and worse long-term welfare outcomes (Berlinski et al. 2009; Moussa and Omoeva 2020).

To increase school enrollment and improve learning outcomes, it is important to develop complementary interventions that target both the demand for education (demand-side), as well as the access to learning materials (supply-side) able to adapt according to the needs of a child. This paper presents experimental evidence on two complementary interventions that use entertainment education and adaptive educational technologies that jointly aimed to address several of the constraints faced by rural communities where traditional social norms govern parental attitudes toward education. The study is a cluster randomized controlled trial designed to experimentally answer the following two questions: *1) Can aspirational videos change parental attitudes toward education and improve enrollment and learning outcomes in traditional societies?* and *2) Can educational technologies (edtech) boost their effectiveness?*

Various interventions have been designed to motivate school enrollment and improve

learning for students. For a systematic review of educational interventions in Africa, see Conn (2017) and Evans and Mendez Acosta (2021). The available evidence suggests that while supply-side interventions can be highly effective, when these involve the provision of inputs alone (e.g., textbooks, flip charts, additional teachers) without the training about how to use them, they have limited impacts on school enrollment, attendance and learning outcomes (Ganimian and Murnane 2016; GEEAP 2020). Demand-promotion interventions such as conditional cash transfer programs have shown a substantial effect on increased school participation (Baird et al. 2019; Schultz 2004), but not on learning outcomes. Teaching at the Right Level (TaRL), which usually involves dividing primary school children into groups based on learning needs rather than age or grade, has been shown to be an effective approach through multiple delivery modes (Banerjee et al. 2007, 2016; Duflo et al. 2011). However, scaling up this approach may present important challenges in low-resource settings, as TaRL models often involve either placing additional teachers in schools or retraining existing teachers to conduct more differentiated instruction. Moreover, in traditional settings where female seclusion or a strong bias for boy education is prevalent, in- and after-school instruction sessions will still need to be complemented with interventions aimed at reshaping parental attitudes towards girls' education. This study aims to bridge the need for supply-side interventions to adapt learning materials to the children's learning needs, while addressing gender-biased social norms that commonly hamper investment returns for girls.

In addition to promoting the demand for education, it is also important for interventions to directly improve learning in low-literacy settings. Input-based interventions that target only the provision of schooling without complementary interventions to improve the quality of learning outcomes are bound to be less effective (Mbiti 2016; Evans and Popova 2016). Starting in the 2000s, a growing wave of education technology or edtech approaches has emerged to address learning gaps in developing countries. The first generation was hardware-focused (e.g., provision of laptops and tablets) and was found to be expensive and ineffective (GEEAP 2020). A handful of recent technology interventions have shown promising results for improving learning outcomes through the facilitation of self-led learning. However, most of these interventions focus on short-term impacts (Angrist et al. 2020a) or take place in urban areas where adverse social norms and parental illiteracy are less binding (Duflo et al. 2011; Banerjee et al. 2016; Muralidharan et al. 2019).

While the evidence on the effectiveness of edtech solutions is still emerging, there is reason for optimism. In particular, edtech may be well-suited for addressing challenges that otherwise require resource-intensive interventions which are very difficult to implement in schools with limited capacity and high pupil-teacher ratios (Rodriguez-Segura

2021). For example, edtech interventions can provide a suitable alternative to classroom-based Teaching at the Right Level interventions by tailoring content to different student learning levels (Banerjee et al. 2007; Muralidharan et al. 2019; Ito et al. 2019). In addition, efforts to create edtech interventions that promote self-led learning through 'gamification', or the application of game elements such as competition and point-scoring to non-game areas such as learning, have shown positive effects on both the ability to keep students engaged and improving their learning outcomes (Araya et al. 2019; Rosas et al. 2003).

This study describes the design and implementation of a trial that was implemented in 128 school catchment areas (SCAs) in northern Nigeria, where communities were randomized into two treatment conditions (32 SCAs per each treatment arm) and one control group (64 SCAs). In all treatment communities, households with 6-9-year-old girls and boys were invited to attend two community screenings of aspirational videos. These videos showed the improved opportunities that education may open in adulthood. In half of the treated communities, and through a lottery conducted at the end of the second screening, 40% of households received a smartphone preloaded with literacy apps. While the videos were greatly based on *vicarious experiences* using relatable characters and stories, the edtech component mostly relied on *enactive mastery experiences* (Bandura 1997) through immediate feedback and experiencing a sense of progress.

Our intent-to-treat (ITT) estimates show that aspirational videos can be an effective tool to increase school enrollment for girls and boys. These videos were also effective at reshaping parental attitudes toward early marriage and increasing their expectations for their children's educational achievement. The mobile learning (m-Learning) add-on increased both literacy and numeracy skills for the target child and older siblings. The combined intervention also increased parental self-efficacious beliefs about their own capability to help their children learn at home, increased parents' aspirations for seeing their children to continue studying in the future and for delaying early transitions into the labor market. Moreover, the combined intervention reduced early childbearing of older siblings.

This study contributes to the literature on educational interventions in communities that are governed by traditional social norms, where the evidence base remains scarce. In addition, it contributes to the economic literature on how to design interventions using social psychology theories and entertainment education formats, aimed at increasing human capital investments by increasing the perceived alternatives parents have for the future of their children, while boosting such impact using digital technologies at home. Finally, this study provides evidence from a middle-income country using mobile-based

technologies to offer a TaRL approach to low-literate populations, contributing to filling the evidence gap pointed out by (GEEAP 2020) about the evidence of the impact of studies using such innovations and their cost-effectiveness.

The rest of this paper is organized as follows. Section 2 provides a background of northern Nigeria, where our study takes place. Section 3 provides an overview of entertainment education. Section 4 describes the interventions and explains the design rationale behind them. Section 5 describes the evaluation design and data collection and Section 6 presents results on main outcomes and mechanisms, spillover effects on other household members, treatment effects by gender, and robustness checks. This section also includes an effectiveness and cost-effectiveness analysis, comparing our learning results with those of other educational interventions. Section 7 concludes.

2 Study Setting

This study took place in Kano and Jigawa states, located in the North West region of Nigeria. Based on model-based estimates, Nigeria has the third largest out-of-school population in the world and the largest in Sub-Saharan Africa with approximately 20 million children out-of-school (UNESCO 2022). This adds up to over 10 percent of the global out-of-school population (Antoninis 2014; Bertoni et al. 2019). The problem is particularly severe in northern Nigeria, which accounts for a disproportionate number of out-of-school children in comparison to the south of the country. In 2018, only 51 percent of the school-age children attended school, versus 84 percent at the national level (The World Bank 2021). Moreover, gender gaps in educational attainment are particularly salient in North West Nigeria, where 64% of 15-49-year-old women have no education in comparison to 38% of men of the same age, according to the Demographic and Health Survey 2018 (NPC and ICF 2019).

In Nigeria, schooling attendance does not necessarily translate into increased learning, and educational indicators for the north substantially lag behind the national average. According to the Reading and Access Research Activity report (2019), in northern Nigeria, less than 3 percent of second grade pupils in government primary schools could read a text in Hausa with 80 percent comprehension or greater (Kano: 2% girls and 3% boys; Jigawa: 0% girls and 3%). One factor that may contribute to lower school enrollment and learning outcomes in the north is the inter-generational persistence of low rates of educational attainment. Overall, in the North West of the country, only 29% of 15-49 year-old women and 59% of men are literate (NPC and ICF 2019). A number of analyses show that parental educational attainment is a strong determinant of educational outcomes among

children in Nigeria (Azomahou and Yitbarek 2016; Razzu and Wambile 2021). Northern Nigeria has some of the poorest educational attainment rates for older adults. Only 40 percent of 30-34-year-olds have ever attended school in the North East and North West zones, compared to 90 percent in the South East and South West regions of the country.

Another factor that may contribute to low school enrollment and deficient learning outcomes is that the region is also characterized by a strong adherence to traditional norms. The formal legal institution of Shari'a law applies in most northern states, covering social, civil, and criminal matters (Suberu 2009; Human Right Watch 2014). This has reinforced social norms that encourage early marriage and childbearing for Muslim adolescents and in the process create additional barriers to education. This is also reflected in attitudes toward education for women and labor force participation of mothers. According to the World Values Survey (2017 -2021), in Nigeria, 42% of respondents believes that university is more important for a boy than for a girl and 41% believes that pre-school children suffer if the mother works. These adverse norms contrast with those observed in other countries where education plays a more prominent role, such as Kenya, where the proportions of the population holding these views are 18% and 23%, respectively.

Further, the emergence of the militant terrorist group Boko Haram, which translates as "Western education is forbidden", creates an additional barrier for school enrollment and attendance. As a result, the predominant forms of schooling available to young children in the north include: 1) Qur'anic schools, which exclusively focus lessons on Islamic theology and religious texts, or 2) Islamiyya schools, which also cover Islamic subjects but follow a more typical school structure with regularized time schedules.[2]

3 Entertainment Education

Entertainment education (EE) encompasses a wide umbrella of interventions that use entertainment media with the goal to positively reshape attitudes and promote behavioral change. The theoretical framework underpinning its narrative format is informed by social cognitive theory and social learning theory, both of which suggest that individuals observe and learn through the success and mistakes of others (Bandura 1971, 2003). By modeling and contextualizing desired behaviors, EE interventions can be more effective than the delivery of information alone. Research finds that EE narratives can promote

[2]Most of these schools do not offer teaching on secular subjects (e.g., science, math), either due to preference or a lack of qualified teachers. Some of these schools, however, have been integrated into the national public school system, with the introduction of secular subjects with full or partial government support (Antoninis 2014).

safer sexual behaviors, greatly driven by program-immersion and character-identification (Banerjee et al. 2019); with meta-analysis evidence highlighting its effectiveness across multiple settings (Orozco-Olvera et al. 2019). This communication approach can potentially influence perceptions of self-efficacy and social norms (Grady et al. 2021). In India, Jensen and Oster (2009) find that the introduction of cable television is associated with improvements in women's status indicators, moving gender attitudes of individuals in rural areas closer to those in urban areas. In Brazil, La Ferrara et al. (2012) find that through the portrayal of small families with empowered female characters, soap operas can affect perceptions of self-efficacy and help reduce fertility rates.

EE interventions can also be effective at communicating updated social norms, and therefore influence behaviors. Social norms significantly mediate what is considered acceptable behavior (Anderson et al. 2015). Experimental research shows that EE interventions that target perceptions of social norms, particularly when implemented in community settings where social groups are collectively present, can have positive effects on behaviors (Blair et al. 2019; Arias 2019; Green et al. 2020). We use EE as the main strategy to change parental attitudes toward education. We describe how we use EE as part of the interventions we use for this study in the next section.

4 The Interventions

In this section, we first describe the components of the interventions followed by their implementation and the design rationale behind them. The implementation and monitoring of the interventions took 5 days. The implementation of screenings was in two days; the delivery of mobile phones and training to parents and target children on the mobile apps was done in one day. Two additional days were used to approach community leaders and have mobilizers visiting households one week and one month after the mobile phones were delivered to the families.

4.1 Aspirational Videos

Design and Implementation

The design of our interventions acknowledged the inter-linkages between education, early marriage and child labor in traditional and low-income communities; hence, the content of our aspirational videos (shown in treatment communities), addressed the three topics having as a core theme the role of education in the future of children. This, in turn, aimed to encourage parents to bring their children to school. To help people imagine and

visualize a future that most parents had not considered before for their children, we decided to use videos instead of other media formats. The entertainment media used in this trial were produced by Impact(Ed), formerly Discovery Learning Alliance, a US-based non-profit media firm that works with ministries of education in developing countries to produce television and mass media initiatives. The study videos are an excerpt from the 10-hour video series "My Better World", an entertaining animation program that also includes real people sharing their experiences in inspiring mini documentaries, often set against a pan-African village life. These everyday role models reflect the animated story's highlighted themes with their personal stories of struggle, courage, and perseverance.

The community screenings of the aspirational videos were implemented in the following way. Community mobilizers invited eligible households, living within a radius of two kilometers of a randomly selected local school, to attend community events that would discuss "relevant issues for parents". When households were visited to deliver such invitations, no mention was made about the main themes of the videos and screening discussions to avoid any potential self-selection of participants against or in favor of education for children. Using data from the listing exercise, the target child (6-9 years old) and both parents were invited to two sessions that would take place in the local school over two subsequent weekends. The implementing NGO was responsible for organizing and staffing the community screenings. NGO female facilitators were trained by the authors about which messages to highlight during the post-screening discussions based on the aspirational video content. To host a maximum of 75 households in the local schools for the screenings and discussions, two parallel screenings (in two different classrooms) took place in most treated communities. The day of the screenings, a registration desk was set up at the entrance to ensure that only invited participants were admitted to the venue. Snacks and a gift bag with a few early-grade literacy booklets were handed to each household head as they completed registration. The community leader welcomed attendees, briefly highlighting the importance of primary education, and explaining the ground rules for participation. The public endorsement of the community leader of the activity led by the NGO facilitators had the purpose of minimizing possible disruptions during the screenings and discussions. The two sessions per community lasted a total of four hours per screening. The aspirational video-clips were on average 5 minutes long and a total of 11 videos were screened in 2 weekends. The videos involved short documentaries and animations of people living in similar social and economic circumstances to those observed in the communities where this study took place. After the video screenings, NGO facilitators led a discussion by asking questions to participants to reinforce the key messages shown in the videos. In total, the screenings and discussions lasted less

than 2.5 hours, including logistical arrangements (e.g., seating participants and filling exit surveys at the end) and a break.[3]

Design Rationale

The aspirational videos aimed to improve parental attitudes toward education and learning by expanding the choice set they had for their children's futures. To do so, the documentary videos showed how education may open up a new life and professional trajectories, such as becoming one day a teacher, a doctor or a business person. The animations complemented this message by acknowledging the challenges of combining school activities with work, as well as of the social pressure for girls and families to comply with the common practice of early marriage. The animations gave practical recommendations to deal with both challenges (e.g., how to make time for homework when the child was part of the family business, how to deal with relatives trying to enforce early marriage when wanting to leave daughters in school, among others). Because of the disparities in school enrollment between girls and boys in northern Nigeria, the selected video-clips had a strong focus on girls being equally capable to benefit from schooling and securing occupations useful for the families and communities in the future (e.g., becoming a doctor, teacher, journalist, among others). To make parents reflect on their own circumstances and on possible future scenarios for their children, videos showed *vicarious experiences* using relatable characters and stories about people living primarily in rural villages in Sub-Saharan Africa. This is consistent with field experiments that have taken place in this region that have shown that increased program immersion and character-identification is greatly attributed to engaging narratives which are effective at improving attitudes and behaviors of viewers (Banerjee et al. 2019; Berg and Zia 2017).

While talking about the situation of a family member can be perceived as intrusive, we used the approach of Bandura (1997) where we encouraged participants to think about a third person. Towards the end of the community screenings, attendees were asked to reflect upon their own circumstances and use the story narratives as a reference point when assessing the pros and cons of new behaviors. Finally, and to reinforce the program's self-efficacy messages, facilitators that led the post-screening discussions were asked to emphasize the point that parents' educational level should not be an obstacle for helping their children to learn.

[3]The authors created a manual to guide facilitators on the key messages of the video-clips, as well as to advise on the time they could spend on each video-clip and discussion of messages. Random checks were done by field coordinators to ensure facilitators followed the agreed protocol.

4.2 M-Learning Add-on

Design and Implementation

In half of the treated communities, and through a lottery conducted at the end of the second screening, a third of the households received a smartphone preloaded with two literacy apps. Both apps can run offline and they are known as: Feed The Monster (FTM) and the Global Digital Library (GDL). The former aims to teach and reinforce foundational literacy skills including sound, letter and word recognition; the latter aims to promote more advanced reading skills through short stories. Both apps were translated into Hausa addressing the common issue found in previous interventions when delivering reading materials in languages that were not the first language used in the communities (Glewwe et al. 2009).[4] Moreover, these apps provide a range of literacy-level material according to the initial skills and progress of the child. In this way, the apps offer the right level of foundational skills needed by the child, and the child progresses according to their pace and level of engagement. This feature of the apps provides a *teach at the right level (TaRL)* approach to learn at home. Previous interventions considering TaRL approaches have primarily been implemented in-school.

The Feed the Monster (FTM) app was produced by Curious Learning, a US-based nonprofit that works with development partners to curate, localize, distribute and evaluate free open-source apps to promote foundational literacy worldwide. Based on game-based learning principles (e.g., for each level, FTM players need to feed the monster as fast and as accurately as possible to gain points and move to the next stage), FTM breaks a language alphabet into small groups of 5-6 letters each. For each group, players must first learn the letters, followed by practicing the combination of letters and their phonetics, to then read a new word with the letters presented in each group. This app has been translated into 50 languages.

The Global Digital Library (GDL) is a global platform developed to increase the availability of high-quality reading resources in under-served languages worldwide. It is part of The Global Book Alliance, an international initiative involving multiple stakeholders working to transform book development, procurement and distribution to ensure that no child is without books. The GDL collects existing high-quality open educational reading resources, and makes them available on web, mobile and for print. GDL books are classified by reading proficiency levels. The first level usually entails a picture matched with one word, serving as a good link with the final levels of the FTM app. The GDL currently

[4]Recent research suggests that learning how to read in one's mother tongue is more effective than learning to read in another language(Brunette et al. 2019). This is a major concern in many developing countries, where many children speak a different language from schools' instruction language.

offers 6,500+ books in 93 languages.

To select the families for the m-Learning add-on, at the end of the second aspirational video screening in half of our treatment communities, a raffle took place to provide 30 households (out of 75, on average) with a low-cost Android device and a solar charger (the market prize for both was around USD 50 in 2019). To avoid selection issues on screening attendance between treatment arms, the prize of the raffle was only announced to the community leader and event attendees after the post-screening discussions of the second session. The selected households were then invited to a 2-hour training delivered three days after the second screening. During the training, one parent (usually the father) and the target child were taught how to use the mobile phone as well as the two literacy apps.

Design Rationale: m-Learning apps

There were three theoretical mechanisms behind the m-Learning component that we anticipated would increase learning outcomes. First, literacy apps can increase *access* to early grade reading materials for home-learning, without restricting the family to specific schedules to attend in-school or after-school instruction sessions, as in Muralidharan et al. (2019). Mobile apps can also increase access to learning materials for young and teenager girls in traditional settings were female seclusion is prevalent and parents may be concerned about social sanctions for sending girls to school. Second, these technology devices could empower non- and low-literate parents to help their children learn at home no matter their educational level (over half of parents in our sample had no education). Parents or other caregivers only needed to supervise children when using the literacy apps (the apps allowed parents to check user progress in terms of levels reached or books opened). Third, and inspired by Bandura's approach of *enactive mastery experiences*, when receiving immediate feedback through the app and experiencing a sense of progress, children may gain confidence in their learning abilities and continue using these apps. In addition, for parents, observing the progress of their children may improve their perceptions about the capabilities of their children to learn. Similarly, through learning-by-playing principles, the "gamification" format of FTM encouraged children and their siblings to play more time.[5] Finally, because users decide the instruction pace, for exam-

[5]This is confirmed by our implementation data. Community mobilizers were asked to visit households after a month of the smartphone provision to answer phone-related questions and to retrieve app usage data. In this first month, the median number of accumulated hours was 6.2 hours for FTM and 2.0 hours for GDL, with the average number of hours more than doubling between the first and fourth week. The authors were not able to collect further usage data during the follow-up data collection because the usage tracker automatically switches off after one month of being active.

ple, the reading level of the GDL book, the apps provide a form of self-teaching at the right level.

5 Evaluation Design

Our study took place between 2018 and 2020 in Kano and Jigawa states, in the North-West region of Nigeria. To focus on demand-side constraints to education, the study sample was restricted to rural communities where local schools had recently received supply-side public investments (e.g., teacher trainings, school materials and small infrastructure projects) and had at least 4 classrooms to potentially enroll new students. School catchment areas were defined by the authors using a 2-kilometer radius around the local school.

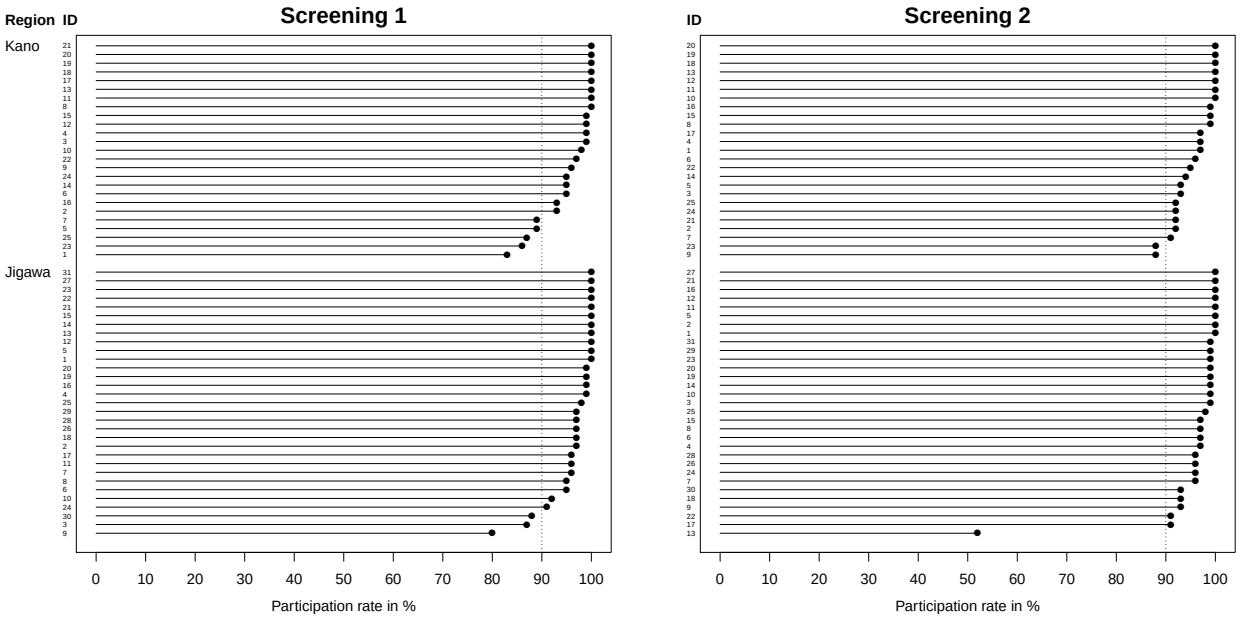
We designed and implemented a cluster randomized controlled trial to evaluate both interventions (aspirational videos and the aspirational videos with m-Learning add-on) where the school catchment area was the unit of randomization.[6] In total, 128 communities were part of this study, where 64 were randomly assigned to receive community video screenings and 64 were assigned to a pure control group. Half of the treated areas were then randomly assigned to receive the m-Learning add-on ($T2$).[7] In all treatment communities, 75 households were invited to the aspirational videos with exception of 8 communities which did not have the minimum amount of eligible families, ranging from 25 to 65 households. In the $T2$ communities eligible for the m-Learning add-on, 30 mobile phones were randomly distributed among the attendees of the community screenings.

To minimize differences in participation rates between the attendees of the screenings-only arm ($T1$) and the screenings-plus-mobile arm ($T2$), we randomly distributed the technology devices at the end of the second screening. The baseline sample includes a total of 9,393 households, with the following breakdown per experimental group: 2,335 households living in 32 communities that received aspirational videos only, 2,345 households living in 32 communities that received aspirational videos and where 40% of them received a smartphone, and 4,713 households living 64 control communities. The take-up

[6] We refer to school catchment areas and communities as synonyms throughout the paper.

[7] Ideally, our design would have maximized statistical power having 42 communities in each experimental arm, however, the decision of assigning 64 communities to treatment and 64 to control was due to the uncertainty regarding the implementation of the m-Learning add-on as it required a separate procurement process for the acquisition and delivery of devices. Because the authors did not have certainty about the exact date the procurement for the m-Learning add-on would be finalized, and the delivery of aspirational videos was linked to a sponsor's deadline, we based our power calculations on the aspirational videos only, splitting by half the sample communities (i.e., 64T and 64C). For the same reason, we also implemented the interventions in the T2 communities after implementing the interventions in T1 communities.

Figure 1: Aspirational Videos: Take-up



(a) Screening 1

(b) Screening 2

rates of the aspirational videos were, on average, above 90% for both screenings. Only 8 school catchment areas had a take-up between 80%-90% for Screening 1 (first weekend) and three of them remained with a take-up below 90% in the Screening 2 (second weekend). Figure 1 presents take-up rates for Screening 1 and 2 for all areas selected for treatment of aspirational videos ($T1$ and $T2$). All families randomly selected for the mobile phones with m-Learning apps ($T2$) picked up the mobile phone (100% take-up) when instructed and assisted the 2-hour-training to explain how to use the phone and the two m-Learning apps.[8]

5.1 Data Collection

Using geo-spatial information provided by the Nigerian Federal Ministry of Education on the location of local schools in Kano and Jigawa states, the authors defined school catchment areas based on a 2-kilometer radius around the closest local school. We then created a sampling frame by producing a list of all households with children between

[8]This evaluation design was pre-registered prior to the baseline data collection at the The American Economic Association's registry for randomized controlled trials `https://www.socialscienceregistry.org/trials/3619`, AEARCTR-0003619, on December 5, 2018. Ethical approval was given by Solutions IRB on December 5, 2018, and renewed with COVID-protocols to minimize infection risks prior to starting the second phase of the follow-up survey on September 3, 2020. Registration number: IRB Registration #: IORG0007116.

the ages of 6-9. We randomly selected eligible households from the list after randomly allocating the school-catchment areas to control and treatment groups. Selected households in treatment and control communities were visited by the survey firm for baseline data collection. A month after the baseline, community mobilizers visited households to invite them to the community screenings. The invitation was for a 6-9 year old and their parents.

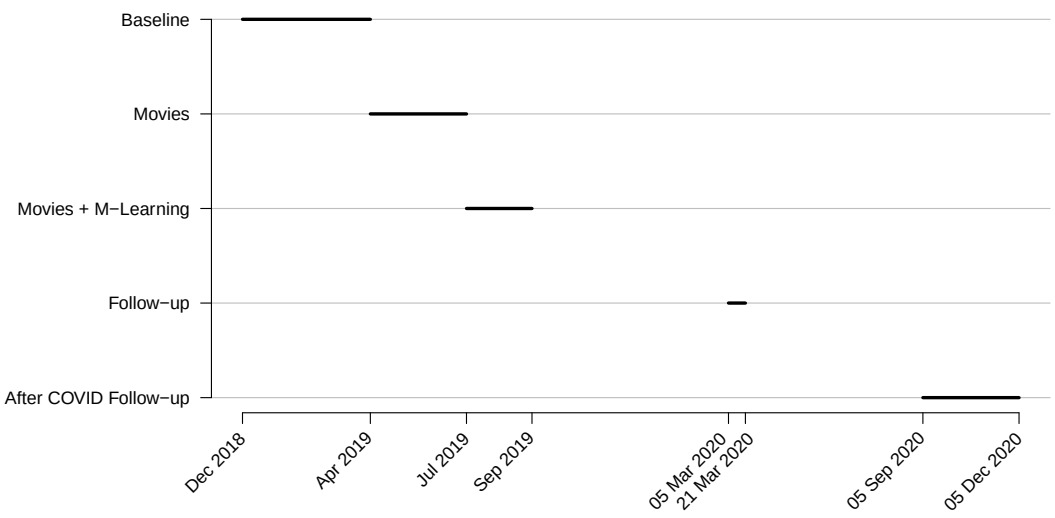
The baseline was collected between December 2018 and April 2019, the interventions were staggered and took place between April-September 2019. Follow-up (endline) data collection started in March 2020, although data collection was stopped three weeks later due to COVID-19 shutdowns. We managed to interview 21% of our trial participants prior to COVID-19 closures. We resumed endline data collection in September 2020 and completed it in December of that year. Figure 2 shows the timeline of the study's data collection activities and of the interventions.

An experienced survey firm (Hanovia Ltd) conducted the trial's data collection. With a support letter from the federal and state ministries of education of Kano and Jiwaga states, the survey firm sought approval from the community leader to provide access to all 128 study communities and to interview selected households to collect listing, baseline and endline data. Baseline data were collected for 9,393 households that were randomly selected from the listing. To maximize the likelihood of finding adults for interviewing and children for testing, the survey firm was asked to visit up to three times selected households, local markets or businesses to find study participants. Attrition rates at follow up were very low: only 3% of main respondents and 5% of target children were not found during endline. Most children who were not found had passed away or migrated to another village or state to attend Almajiranci (Islamic school), a religious and non-formal system of education practiced in northern Nigeria.

5.1.1 Main outcome measures

The main outcomes are the percent of children not attending school and learning outcomes (literacy and numeracy test scores) as pre-registered in the AEA registry of randomized control trials, see footnote 8. For both baseline and follow-up data collection we interviewed the household head and tested on literacy/numeracy skills a 6-9 year old child (referred as 'target child', as this child was invited to the community screenings and targeted by the interventions' messaging) and a 6-12 year old sibling (referred as 'sibling'). One child per age group was randomly selected at baseline if the household had more than one child in each age category. Assignment was stratified by gender of the target child. To analyze the mechanisms through which our interventions may

Figure 2: Timeline: Baseline and Endline Data Collection



have worked through, we also measured parental aspirations and expectations for their children's futures, parental self-efficacy beliefs, perception of social norms and parental time allocation for home-learning activities. Finally, to study potential spillovers on siblings and other household members, the baseline and follow-up surveys collected data on school enrollment, learning outcomes, early parenthood, and early insertion in the labor market.

To measure the proportion of out-of-school children, we use two questions depending on whether follow-up data was collected before or after COVID-19 closures. Prior to March 2020, we asked parents if the target child was attending formal schooling (this represented 21% of the overall sample). For the remaining participants, and once the follow-up data collection resumed in September 2020, we asked a retrospective question about the child attending formal education between January-March 2020. See Section 5.5 for more information about the implications of COVID-19 and our robustness checks. [9]

[9]Because our main outcome on out-of-school children is based on parents' responses, our measure might be subject to social desirability biases or experimenter-demand effects. Data checks of parental responses with school registry records were not possible in our study due to implementation and cost constraints. To address this potential issue, we collected a 'wisdom of crowd' measure that does not rely on direct information about the target child, and therefore it is less sensitive to such biases. The 'wisdom of crowds' measure aggregated the responses to the following question: "Now I would like to ask you about girls/boys like [Child's name]. Think now about 10 children who are the same age, gender and intelligence as [Child's name]. These 10 children have parents with the same level of literacy, education and socio-demographic characteristics as you and your spouse/partner. How many of these 10 children do you think will attend school in January 2021? If you do not know the answer, try to give your best guess." This question was adjusted for September 2020 if the household was interviewed in March 2020 and for January 2021 if the household was interviewed in September-December 2020, as a result of the pause in data collec-

For literacy and numeracy outcomes, we aggregated all levels of literacy and numeracy skills and constructed a normalized score measure. This measure consists of normalizing literacy and numeracy scores using the mean and standard deviation for the control in $baseline$. This approach allows comparing our results with other studies targeting primary school age children with edtech strategies such as Banerjee et al. (2007), Muralidharan et al. (2019), among others. Our main analysis also reports the proportion of children providing 'no correct' answer in letter recognition, a measure of basic literacy foundations commonly used in the education literature.

To understand what competency levels the interventions may have improved, we assessed the target child and sibling using *Early Grade Reading Assessment* (EGRA, literacy) and Early Grade Mathematics Assessment (EGMA, numeracy) in Hausa. We used each level of competency to construct an aggregate of learning measures. All levels are reported in Table 1.[10]

5.1.2 Mechanism measures

As described in Section 4, the videos aimed to help parents visualize a future that many of them had not considered before for their children. This involved increasing the aspirations of treated parents by expanding the mental set of choices they had for their children if they were educated. To measure how much our interventions changed the way parents perceived the future of their children, we collected data on parental aspirations and expectations for the target child. For aspirations, we asked parents about what they desired their children to become or to do. For instance, we use questions such as *What is the highest level of education you would like your child to complete?* and *At what age would you like your child to get married?*. To measure expectations, we also asked parents about the chances that certain outcomes would happen in the child's future. For example, we use questions such as *how likely do you think it is that your child will finish secondary school?* and *how likely do you think it is that your child will get married when your child turns 15?*

To measure parental self-efficacy beliefs about themselves and their children, we asked

tion for COVID reasons. Using the 'wisdom of crowds' aggregate, a measure that is less sensitive to social desirability biases as it is not about the target child in the household, finds that 79.8% (20.2% out-of-school) of the children are expected to go to school in the next available opening of school admissions. Our aggregated measure constructed from the direct question "Is [Child's name] currently attending formal school?" gives a proportion of children attending school of 80.9% (19.1% out-of-school). This reveals that our own measure of school enrollment (using a measure of the target child currently attending school) generates the same aggregate information as the measure that is less sensitive to experimenter-demand effects and social desirability biases.

[10]EGRA non-word reading assessment was not used in the aggregate measure of literacy skills because of the small number of observations for this assessment.

parents (i) how capable they believed they were for helping the target child acquire literacy, numeracy and communication skills; and (ii) their perception of the target child's ability to learn. To complement our analysis on the possible mechanisms that may have boosted the impact of the interventions, we also explore the social norms perceived by the main respondent regarding schooling, early marriage and early insertion into the labor market, as well as the time allocated by the parent on school work and learning for the target child, and activities done by the child.

5.1.3 Spillover measures

To measure spillover effects on school enrollment, we created a binary measure based on a household roster where we identified 6-12 year siblings attending formal school in March 2020 (or between January and March 2020 if the household was interviewed once we resumed data collection in September 2020). To measure learning spillovers on 6-12 year old siblings, we used the same testing strategy used for the target child (EGRA and EGMA tests) but different exercises. For early parenthood, we created two measures: a) early parenthood for family members still living in the household who were under the age of 18 at baseline and b) a measure using the same criteria (a) but adding members who left home as a result of marriage or childbearing. Finally, early insertion in the labor market was measured through the household roster for all family members between 5-20 years old who reported employment in the past month. We also analyzed spillovers on school enrollment and employment for the siblings that were subject to literacy and numeracy tests. The age of all household members is based on the age reported at baseline.

5.1.4 Study Sample

The top half of Table 1 shows descriptive statistics of our sample households at baseline. On average, mothers were 12 years younger than fathers (31 vs 43 years of age).[11] Fifty-three percent of household heads had no formal education, and 20% and 14% respectively had either some primary or secondary education. Household heads are primarily the father or grandfather living in the same household as the target child. Households have an average of eight members and most faced food insecurity issues. Approximately three-quarters reported not always having enough food 7 days prior to the interview and 10 percent reported facing food scarcity sometimes or frequently.

[11] The maximum value for mothers' and fathers' age is above 95 years old. Although these values are not reasonable for biological parents, when looking at the distribution of ages, we see a stable 1% of observations reporting ages above 60 years old, we decided to leave them for the analysis without any manipulation as this may reflect non-biological parents who took the role as parents.

The bottom half of Table 1 displays characteristics and learning statistics of the target child at baseline. Fifty-one percent of target children are female and the average age is 7 years old. Only 17% of target children were not enrolled in school, which is substantially below the average of northern Nigeria where around half of children are not enrolled in primary school (Nigerian DHS 2019). Similar to previous studies implemented in northern Nigeria (Reading and Access Research Activity report 2019), literacy and numeracy levels in our study sample were extremely low. Using modules from the Early Grade Reading Assessment (EGRA), the average number of correct answers in letter recognition was 0.35 letters out of 100. For this first assessment, if the child was not able to give at least one correct answer for the first 10 letters, the exercise was discontinued and the assessment instrument moved on to oral passage reading. Only 1.2% of target children passed from letter recognition to non-word reading. In listening comprehension, where children were asked questions about a passage that was read by an enumerator, children were better equipped, with the average number of correct answers reaching almost 3 out of 5 questions. For letter dictation, on average, approximately one out of five letters was correct.

Considering the overall low literacy rates in northern Nigeria, we also collected Uwezo literacy assessments as their implementation by non-pedagogical experts is simpler than EGRA. This is a household-based initiative that aims to improve literacy and numeracy skills among children in Kenya, Tanzania and Uganda. A local translator used Uwezo assessments and translated into Hausa.[12] The main difference between EGRA and Uwezo is that the number of letters and words tested in Uwezo is fewer, which might encourage the child to provide more responses; and unlike EGRA, the Uwezo assessment is not timed. In our sample, children performed better for the competence of letter recognition using Uwezo assessments, however, the patterns across literacy levels found in EGRA and Uwezo are very similar. Thus, this paper reports EGRA scores to be comparable with other education technology (edtech) studies.

Because of the brain activity that the m-Learning add-on may encourage while playing and learning, we also assess potential spillovers to numeracy outcomes for the target child and sibling (the printed materials provided at the community screenings and the m-Learning apps only included literacy lessons), we used modules from the Early Grade Mathematics Assessment (EGMA) to test both children. Overall, while we observe children doing better at recognizing numbers than letters, the baseline data shows that numeracy skills are very low. On average, 0.43 numbers are recognized out of 20 and 1.6

[12]More details about Uwezo assessments can be found at `http://nada.uis.unesco.org/nada/en/index.php/catalogue/179`.

numbers out of 10 are correctly discriminated – when children need to identify the highest number from a pair of numbers. For the recognition of missing numbers, children also perform very poorly (on average 0.20 correct answers out of 10). And for more advanced numeracy skills, addition and subtraction, children score, on average, 0.20-0.30 correct answers out of 20 numbers.

Table 1: Sample Descriptive Statistics at Baseline

	Mean	St.Dev	Min	Max	N
		Household characteristics			
Age of target child's father	43.88	10.72	15	97	9,029
Age of target child's mother	31.45	8.11	15	98	8,907
Hh education: none	0.53	0.50	0	1	9,393
Hh education: Islamiyya / other	0.06	0.23	0	1	9,393
Hh education: any primary	0.20	0.40	0	1	9,393
Hh education: any secondary	0.14	0.35	0	1	9,393
Hh education: higher than secondary	0.07	0.26	0	1	9,393
Household size	8.61	3.95	2	37	9,393
No. of rooms	2.88	1.60	1	20	9,393
Food: always eat enough	0.18	0.38	0	1	9,393
Food: not always eat enough	0.72	0.45	0	1	9,393
Food: sometimes/frequently not eat enough	0.10	0.30	0	1	9,393
		Target Child			
Female (=1)	0.51	0.50	0	1	9,393
Age	7.31	1.04	6	9	9,393
Out-of-school	0.17	0.38	0	1	9,393
EGRA letter recognition	0.35	3.63	0	92	9,285
EGRA nonword reading	3.98	6.56	0	29	116
EGRA oral passage reading	11.09	1.75	0	56	9,285
EGRA reading comprehension	0.02	0.27	0	5	9,285
EGRA listening comprehension	2.88	1.79	0	5	9,285
EGRA letter dictation	0.80	1.23	0	5	9,285
EGRA word dictation (correct word)	0.05	0.40	0	5	9,285
EGRA word dictation (all letters correct)	0.06	0.42	0	5	9,285
EGMA number recognition	0.43	1.87	0	20	9,273
EGMA number discrimination	1.60	2.49	0	10	9,285
EGMA missing number	0.20	0.85	0	10	9,285
EGMA addition	0.30	1.60	0	20	9,285
EGMA subtraction	0.20	1.35	0	20	9,285

5.2 Empirical Strategy

Our main analysis is based on intent-to-treat (ITT) estimates for the target child where we use the following empirical specification:

$$Y_{1,ic} = \alpha + \beta_1 T + \gamma Y_{0,ic} + \boldsymbol{X}_{t,ic}\Omega + \epsilon_{1,ic}$$

$Y_{1,ic}$ corresponds to the outcome of interest for the target child i living in catchment area c, collected in endline (Wave 1); T is a treatment binary variable where $T = T1 + T2$; $Y_{0,ic}$ is the outcome of interest collected in baseline (Wave 0). To analyze the effect of each intervention, we distinguish between ($T1$) and ($T2$) in a separate regression. We use linear probability models (LPM) for the percentage of out-of-school children and OLS for all z-scores of literacy and numeracy skills. In all regressions the standard errors are clustered at the school catchment area level, our unit of randomization. To improve the precision of our estimates, all regressions include as covariates ($\boldsymbol{X}_{t,ic}$): age of target child, gender of target child, education of household head (at baseline), availability of food (at baseline) as described in Table 1 and enumerator fixed effects.

For the analysis of mechanisms and spillovers, we used the same empirical specification as in the main outcomes analysis, controlling for the same covariates and clustering standard errors at the school catchment area. Similarly, our analysis of treatment effects by gender follows the same specifications, but adding the interactions of the treatment binary variables with a binary indicator for female.

6 Results

We first explore the statistical balance between treatment groups and control for a list of socio-demographic characteristics and our main outcomes (school enrollment and learning skills). We observe in Appendix Table A1 a high statistical balance when comparing treatment groups and control, for each treatment and when both treatment arms are combined. Statistical balance at baseline combined with low attrition rates at follow up (between 3% for main respondent and 5% for target child) lend support to the causal interpretation of our impact estimates.

Our ITT estimates are presented in the following order. First, we discuss the overall impact on the percentage of out-of-school children and learning outcomes when combining samples from both treatment arms: aspirational videos only ($T1$) and aspirational videos with the m-Learning add-on ($T2$). In other words, comparing any treatment to the control group. To assess which component may be driving our results, we then assess the

separate effects that each treatment arm had on these outcomes. Second, to understand the underlying mechanisms that may be driving our results, we analyze the channels of (1) parental aspirations and expectations regarding school achievement, early marriage, early childbearing and early employment; (2) parental perceptions regarding their capability for helping their children learn and their perception of their children's capability to learn; (3) social norms perceived by the parent and (4) time allocation of parents and children on learning-related activities. Third, we discuss spillover effects on other family members for learning, early pregnancy and early transition to the labor market. Finally, we study treatment effects by gender of the target child and future aspirations and expectations of parents for their daughters and sons. This section also provides various robustness checks to assess the stability of our main conclusions given the disruption of follow-up data collection as a result of COVID. We close this section by presenting results of a cost-effectiveness exercise where we compare our results on learning outcomes of the combined intervention ($T2$) with respect to other educational investments.

6.1 Out-of-School Children and Learning Outcomes

Table 2 shows the intervention's impacts on the percentage of children not enrolled in school. The first column presents the overall treatment effect, without distinguishing between $T1$ and $T2$; and the second column separates results for each treatment. Compared to the control group, children living in treatment communities are 6.3 percentage points (42.6 percent) less likely to be out of school. The m-Learning add-on did not have an additional impact on school enrollment, the coefficients of both treatments are practically the same (T1 vs T2: *p-value* = 0.899). Although the evidence suggests the quality of teachers' training in rural Nigeria is extremely low, our findings suggest that despite this supply constraint, the mobile-based component did not substitute learning at school for learning at home.

On learning outcomes, Table 3 shows that children in treatment communities became less likely to have a 'zero' score (no correct answer) in letter recognition – the first level of literacy. As demonstrated by column 2, this effect was driven by communities that received the smartphone with m-Learning apps. While $T2$ decreased by 18.6 percentage points (42.5 percent) the number of children that tested 'zero' in this literacy level, the aspirational videos alone did not result in better learning outcomes despite of the increase in school enrollment observed in the households allocated to this intervention.[13] Looking into the overall learning z-scores, $T2$ increased literacy and numeracy skills by 0.46 and

[13]This is further discussed in the section 6.5 in the robustness checks.

Table 2: Out-of-School Children (Formal School)

	(1)	(2)
Any Treatment (T1 or T2)	-0.063***	
	(0.013)	
Videos (T1)		-0.063***
		(0.014)
Videos + Mobile apps (T2)		-0.064***
		(0.014)
Observations	9,201	9,201
R-squared	0.239	0.239
Covariates	YES	YES
Enumerator FE	YES	YES
Control Mean in F-up	0.148	0.148
T1 vs T2 (p-value)		0.899

Note: Linear probability models with clustered standard errors at the school catchment area level. The covariates considered in these regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

0.63σ respectively. [14] When comparing both components, we confirm these effects were driven by the combined intervention of aspirational videos plus the m-Learning add-on (T1 vs T2: *p-value*=0.000). The aspirational videos alone did not have significant impacts on learning skills. [15]

Tables A2 and A3 in Appendix show results for disaggregated levels of competences assessed by EGRA and EGMA modules. The combined intervention ($T2$) improved scores for almost all literacy and numeracy competence levels, with statistically significant increases ranging from 0.16 and 0.67σ. The largest impacts are observed in foundational skills for both literacy and numeracy. For literacy, the largest impact is observed in 'letter recognition' and 'letter dictation', where children living in $T2$ school catchment areas scored 0.42 and 0.41σ higher than the control group, respectively. This is followed by 'listening comprehension' (0.23σ), 'oral passage reading' (0.18σ) and 'word dictation' (0.16σ).

[14] Using data from the Uwezo literacy assessments, which are simpler tests designed for low-literacy populations, we find qualitatively similar results in terms of literacy gains. These results are not shown in the paper.

[15] Our learning outcomes have been standardized using the mean and the standard deviation of the control group at baseline as in Banerjee et al. (2007) and Muralidharan et al. (2019) to facilitate the comparison of treatment effects across studies. See cost-effectiveness analysis in Section 6.6.

In $T2$ communities, the largest numeracy impacts were observed in 'number recognition' (0.67σ), 'missing number' (0.47σ) and 'number discrimination' ($0.43\ \sigma$) when compared with the control group. These impacts are followed by addition and subtraction (both indicators with gains of 0.32σ). For all these indicators, the difference between $T2$ and $T1$ is significantly different from zero.

Table 3: Treatment Effects on Literacy and Numeracy Test Scores

	Literacy Skills				*Numeracy Skills*	
	Prop. Zero		Z-score		Z-score	
	(1)	(2)	(3)	(4)	(5)	(6)
Any Treatment (T1 or T2)	-0.077**		0.187*		0.243**	
	(0.036)		(0.106)		(0.116)	
Videos (T1)		0.024		-0.071		-0.117
		(0.043)		(0.108)		(0.132)
Videos + Mobile apps (T2)		-0.186***		0.462***		0.628***
		(0.034)		(0.132)		(0.134)
Observations	8,928	8,928	8,928	8,928	8,928	8,928
R-squared	0.163	0.183	0.248	0.253	0.264	0.280
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Control Mean in Baseline	0.976	0.976	0.000	0.000	0.000	0.000
Control Mean in F-up	0.438	0.438	-0.713	-0.713	1.063	1.063
T1 vs T2 (p-value)		0.000		0.000		0.000

Note: Column 1 and 2 show the results of linear probability models of the proportion of children scoring "zero" in letter recognition. Column 3-4 show z-scores of literacy skills and Column 5-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Treatment effects for $T2$ communities should be read as lower-bound estimates given that only 40% of households attending the aspirational videos received a smartphone with the m-Learning apps. Thus, our ITT estimates include both participants receiving and not receiving the mobile phone.

6.2 Mechanisms

In this section we discuss some of the mechanisms that may explain the changes in our main outcomes. We focus on three psychological measures such as parental aspirations, expectations, and self-efficacy beliefs, as well as on social norms and time allocation.

6.2.1 Parental aspirations and expectations

The interventions were successful at expanding the choice set parents had for their children's futures. Appendix Table A4 shows that parents living in treatment communities are more likely to aspire to have the target child in school at the age of 15 and 18 than the control group participants. Parents are also less likely to want their children to be married at those ages. When we disaggregate the treatment effects, we observe that the aspirational movies with m-Learning apps primarily changed aspirations for school, work and apprenticeships ($T2$). Increasing aspirations for school by 5.5 percentage points (pp.) and decreasing aspirations for work and apprenticeships by 1 pp. While $T2$ was effective at increasing aspirations for activities that are related to the use of skills, the treatment of only aspirational videos $T1$ was primarily effective at decreasing aspirations of seeing children getting married at early ages. This is confirmed when using a continuous measure of the age parents aspired to see their children getting married and having their first child. Appendix Table A5 shows that the interventions increased the age parents aspired for their child (boy or girl) to get married and have a child by 0.44 and 0.39 years, respectively.

When looking at the chances of observing certain outcomes in the future (parental expectations), Table A6 shows increases in expectations ranging from 2.1 to 6.8 percentage points (2 to 11 percent relative to the control mean) in the likelihood for their children of completing primary, secondary and high school compared to the control group. These effects are primarily driven by the aspirational videos as we do not find significant differences between $T1$ and $T2$. When looking at parental expectations about the target child's chances of entering the labor market and getting married in the near future, we observe that only the intervention that involved the acquisition of skills ($T2$) decreased the expectations of observing both outcomes by 20 and 68 percent, respectively.

6.2.2 Parental and child self-efficacy

Because our interventions aimed to increase the feeling of empowerment of parents, regardless of their level of education, we also collected data on how parents perceived their own capabilities to help the target child learn, as well as their own perception about

their children's capabilities to learn. The intervention of videos with an m-Learning add-on ($T2$) improved parents self-efficacy beliefs about themselves and the target child. Appendix Table A7 shows that parents who were allocated to $T2$ improved their self-efficacy beliefs in their capacity to help their child learn (increases between 3.9 and 4.4 percentage points – 4 and 5 percent, respectively) and to improve learning by setting a time and space to study (4.1 percentage points– 4 percent). The combined intervention $T2$ also decreased the perception of parents that their level of education is an obstacle for helping their child learn or improve their skills (an increase of 3.7 percentage points representing an increase of 29.1 percent with respect to the control group). Although the aspirational movies (and post-screening discussions) highlighted the role of parents in the education of their children, regardless of their level of education, we do not find significant impact of $T1$ on self-efficacy beliefs. Differences between treatment arms are statistically significant in most of the cases at the 1% significance level. Considering that more than half of household heads has no education in our sample, this finding suggests that the m-Learning component empowered parents in their perception about being able to help their children's learning pathways, regardless of their level of education.

Table A8 in Appendix also shows that parents in treatment communities ($T1$ and $T2$) improved their perception about their children's capabilities to learn literacy and numeracy skills, including the way they rate their children's ability in comparison to other children with similar family backgrounds. Improvements range from 2.3 and 3.3 percentage points (which represent on average increases of around 5 percent). While the coefficients for the combined intervention $T2$ are statistically significant and greater in magnitude than those for $T1$, we do not find significant differences between both treatment arms on parental perceptions about their children's capabilities. Overall, the perceptions of parents about their capability to help learn did change as a result of $T2$. And being allocated to either $T1$ or $T2$ did change parental perceptions about their children's capabilities to learn.

6.2.3 Setting time and space for study

We also look at actual changes in behavior related to school and learning outcomes. As part of the aspirational videos, the facilitators who led the screenings and discussions were asked to reinforce key messages from the videos (e.g., parents can help children learn through reading, homework, make time for school activities, among others). In this section, we analyze the self-reported time allocation of the parent who was interviewed (in most of the cases the father) and target child.

As expected, Table A9 shows that those families allocated to any of the two treatments

are more likely to read to the target child, help the child learn, make time for school and learning activities. Overall, $T2$ was more effective at encouraging parents to make time for school work or learning activities than $T1$ (*p-value*=0.022), teach names of objects (*p-value*=0.008) and bringing the child to school (*p-value*=0.099). For the target child, we only observe changes in activities related to school attendance (increase by 0.8 hours out of 12) and working outside home to earn money for T2 (decrease by 0.10 hours out of 12), see Appendix Table A10.

6.2.4 Social norms

Community screenings included the discussion of the videos' content immediately after their screening ended. At the start of these community events, the community leader welcomed the female NGO facilitator who would lead the overall session, and explained the main reason why the facilitator and her team were visiting the community (i.e., to talk about education for their children) and legitimize the message of 'talking about education for everyone' in the community. After the screening ended, the facilitator reinforced key messages of the screenings and encouraged a discussion among participants. On average, 75 households attended the community screenings, offering attendees the opportunity to learn about the views and reactions of other community members related to education, early marriage and early insertion into the labor market.

In this section, we analyze whether the interventions made a difference in parental attitudes towards attending school, getting married and starting to work at different ages (6, 9, 12 and 15 years old). We asked parents (main respondents) to think about specific hypothetical scenarios and share with us their opinion, as well as to provide their best guess about what they think most people in their village would answer if they were asked the same questions. For the elicitation of parental attitudes, we created an index of two types of questions: a) whether they considered appropriate, inappropriate or were neutral about seeing a boy/girl attending school/getting married/working outside home at a specific age and b) whether they think there is a positive/negative/no effect on marriage if a boy/girl attends school at a specific age. For the elicitation of the perceived social norm in the village, we adapted question (b) to obtain the respondents' beliefs about what most people would respond if they are asked about the effect on future marriage if a girl/boy attends school at a specific age.

Table A11 shows that $T1$ was effective at increasing attitudes in favor of having boys and girls attending school, and also at decreasing attitudes in favor of early entry to the labor market for boys. Overall, the aspirational videos were effective at changing social norms about school attendance and entry to the labor market. For early marriage, table

A12 shows that $T2$ reduced the main respondents' attitudes (and their perceptions about the village inhabitants') in favor of early marriage for girls.

6.3 Spillovers on Other Family Members

Table 4 shows that the combined intervention $T2$ decreased early childbearing by 1.4 percentage points, a decrease of 13 percent. When considering the family members that also left home between baseline and follow-up as a result of marriage or childbearing responsibilities, the coefficient shows a decrease in childbearing and marriage of 3.9 percentage points, a decrease of 15 percent. While the combined intervention did not have spillover effects on school attendance for 6-12 year old siblings, it did decrease early insertion into the labor market for 5-20 year old siblings. This decrease was in the order of 2.4 percentage points, a decrease of 14 percent. The impact of $T2$ is significantly different from T1 for early marriage, early childbearing, and work outcomes. Similar to the results presented in Table 2, there is no significant difference between $T1$ and $T2$ in school attendance.

Table 4: Spillovers: Early Childbearing Marriage, School Attendance and Work

	Early Childbearing & Marriage (<18 yo)				School Attendance (6-12 yo)		Work in past month (5-20 yo)	
	Becoming a Parent (Members still) at home)		B. Parent + Married (Adding members who left)					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any Treatment (T1 or T2)	-0.003		-0.004		0.016		-0.005	
	(0.007)		(0.017)		(0.016)		(0.012)	
Videos (T1)		0.008		0.029		0.018		0.012
		(0.009)		(0.023)		(0.018)		(0.015)
Videos + Mobile apps (T2)		-0.014*		-0.039**		0.013		-0.024*
		(0.008)		(0.016)		(0.018)		(0.013)
Observations	9,201	9,201	9,201	9,201	5,876	5,876	8,121	8,121
R-squared	0.285	0.285	0.116	0.119	0.267	0.267	0.110	0.111
Covariates	YES	YES	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES	YES	YES
Control Mean in F-up	0.108	0.108	0.263	0.263	0.741	0.741	0.172	0.172
T1 vs T2 (p-value)		0.025		0.002		0.784		0.022

Note: Column 1 and 2 show the results of linear probability models of the proportion of adolescents (under 18 at baseline) having a child. Column 3 and 4 use as dependent variable the same as in Column 1, adding the adolescents that left home at endline as a result of childbearing responsibilities or marriage. Column 5-8 show OLS regressions using as dependent variable the proportion of children 6-12 who are attending school, excluding the target child (Column 5 and 6) and the proportion of members 5-20 who worked in the past month prior to data collection. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

The combined intervention of aspirational videos with the m-Learning add-on also had important spillover effects on learning outcomes of 6-12 year old siblings – defined according to their age at baseline. Table 5 shows that for those siblings in $T2$ communities, the proportion of zeros in letter recognition decreased by 13.8 percentage points, a 29 percent decrease. We also observe the overall literacy and numeracy scores for siblings improving by 0.34 and 0.47σ, respectively. These learning impacts are almost as large as those observed for the target child (i.e., 0.46 and 0.63σ). Interestingly, in $T1$ communities, we observe negative spillovers in literacy and numeracy scores in the order of -0.19 and -0.26σ, respectively. These negative impacts are mainly driven by households with older siblings (13-18 years old) living in both types of treated communities, see Table A13. Older siblings living in $T1$ communities are less likely to be enrolled at school and more likely to have worked in the past month, see Table A14 of Appendix. When looking at the activities done with the siblings (not shown in this manuscript), there is no evidence of parents reallocating more time resources to younger siblings than to older siblings, however, our analysis on the siblings' school enrollment and work suggest that the declined in the siblings' learning skills may have been driven by the shift from school to work.

Table 5: Spillovers: Literacy and Numeracy Skills of Siblings (6-12 year olds)

	Literacy Skills				*Numeracy Skills*	
	Prop. Zero		Z-score		Z-score	
	(1)	(2)	(3)	(4)	(5)	(6)
Any Treatment (T1 or T2)	-0.050		0.064		0.092	
	(0.038)		(0.116)		(0.122)	
Videos (T1)		0.032		-0.190*		-0.257**
		(0.046)		(0.115)		(0.123)
Videos + Mobile apps (T2)		-0.138***		0.340**		0.471***
		(0.039)		(0.145)		(0.141)
Observations	5,268	5,268	5,268	5,268	5,268	5,268
R-squared	0.117	0.130	0.198	0.204	0.227	0.243
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Baseline Mean	0.942	0.942	0.000	0.000	0.000	0.000
Control Mean in F-up	0.476	0.476	0.932	0.932	1.127	1.127
T1 vs T2 (p-value)		0.000		0.000		0.000

Note: Siblings' age is defined according to their age at baseline. Column 1 and 2 show the results of linear probability models of the proportion of children scoring "zero" in letter recognition. Column 3-4 show z-scores of literacy skills and Column 5-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

When we separately identify the households that received the mobile phone with m-Learning apps and those who did not receive the device living in $T2$ communities, we observe large spillovers on target children that did not receive the device. Table 6 shows that treatment effects on foundational skills (column 2) and numeracy (column 6) are larger for target children receiving the device. Moreover, the effects on literacy and numeracy skills are significant for target children not receiving the device. Treatment effects on foundational literacy and numeracy skills on target children without mobile phone, living in $T2$ communities, are approximately 75% as large as the effect observed on target children directly benefiting from the mobile phone with m-Learning apps. When analyzing siblings within $T2$ communities, the spillovers effects are as large as those observed on siblings living in households that received the mobile phone, see Table A15.

Table 6: Spillovers: Literacy and Numeracy Skills of target children in $T2$ communities

	Literacy Skills				*Numeracy Skills*	
	Prop. Zero		Z-score		Z-score	
	(1)	(2)	(3)	(4)	(5)	(6)
Any Treatment (T1 or T2)	-0.077**		0.187*		0.243**	
	(0.036)		(0.106)		(0.116)	
Videos (T1)		0.024		-0.071		-0.117
		(0.043)		(0.108)		(0.132)
Videos + Mobile (T2)		-0.214***		0.550***		0.730***
		(0.036)		(0.136)		(0.139)
Videos + No Mobile (T2)		-0.166***		0.402***		0.558***
		(0.034)		(0.147)		(0.139)
Observations	8,928	8,928	8,928	8,928	8,928	8,928
R-squared	0.163	0.184	0.248	0.253	0.264	0.280
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Baseline Mean	0.976	0.976	0.000	0.000	0.000	0.000
Control Mean in F-up	0.438	0.438	-0.713	-0.713	1.063	1.063
T2 mob. vs T2 no mob. (p-value)		0.017		0.164		0.030

Note: Column 1 and 2 show the results of linear probability models of the proportion of children scoring "zero" in letter recognition. Column 3-4 show z-scores of literacy skills and Column 5-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Overall, the spillovers observed in $T2$ communities are encouraging as they reveal that the m-Learning add-on not only impacted learning outcomes of the target child, but also impacted learning and behavioural outcomes of siblings. Moreover, within $T2$ communities, we observe large spillovers on target children and siblings who did not receive the mobile phone with m-Learning apps.

6.4 Heterogeneous Effects by Gender and Age

Appendix Tables A16 and A17 present results on the main outcomes of this study interacting treatment variables with a female binary variable. Overall, the interventions where equally effective for girls and boys. When exploring heterogeneous effects by age and gender, A18 shows no heterogeneous effects by age for boys – treatment effects are the same for children between 6 to 9 years old based on their age at baseline. For girls, we observe a slight variation; all children are similarly affected by $T2$ (Videos and Mobile Apps), with the exception of 7-year-old girls presenting a smaller impact of $T2$ on learning skills.

Because of the content of our interventions, we expected differences in the mechanisms that made our main outcomes improve. Appendix Table A19 present treatment effects on parental aspirations for daughters and sons, separately. Consistent with the traditional setting of our trial, we observe parents in the control group to have substantially lower aspirations and expectations for their daughters for school enrollment and early marriage compared to their sons. For example, while 27.1% and 54.7% of parents aspire for their daughters to be married by the age of 15 and 18 respectively, less than 0.5% of parents wish their sons to be married at those ages.

With respect to the control group, parents in treatment communities became 6.0 and 7.8 percentage points more likely to want their daughters to be in school by the age of 15 and 18 respectively (8.5% and 19% increases). On the other hand, this effect was smaller for sons and was restricted to when they reached the age of 18 (2.2 percentage points, a 2.% increase). This is largely explained by the high percentage of parents that would like to see their sons to be enrolled in school at the age of 15 and 18 (95.5% and 90% in control group, respectively). Similarly, while parents in treated communities were less likely to wish for their daughters to be married by the age of 15 (6.3 percentage points, 23.2% decrease) and by the age 18 (7.3 percentage points, 13.3% decrease); we observe no effects on parental marital aspirations for their sons. This is largely explained not only by the social practice of seeing boys getting married much later than girls in Northern Nigeria, but also because the video content was primarily focused on the inter-linkages of school, marriage and early insertion into the labor market for girls. We also observe that $T2$ had small effects on parents' wishes to delay their daughters and sons' early entrance into the labor market, including a decrease in their aspirations for their sons to be working under an apprenticeship.

For girls and boys, Appendix Table A20 shows similar increases regarding parental expectations related to attending school the following year (January/February 2021) and similar effects on expectations for completing primary and lower secondary school. The increase in expectations for completing high school was substantially higher for girls (6.6 percentage points, 15.7% increase) than for boys (2.8 percentage points, 3.6% increase). Similar to parental aspirations, this difference in impacts is largely explained by the difference in expectations for girls and boys in the control group (42% for girls vs 78% for boys). For the immediate future question of the girl getting married the following year, we observe the combined intervention $T2$ having a substantial impact on this outcome. Parents in this treatment arm became 39.9 percentage points less likely to expect this outcome, a decrease of 84.9 percent with respect to the control group. Finally, Appendix Table A21 shows that for both main girls and boys, only $T2$ was effective for increasing

parental self-efficacy beliefs and creating time and space for learning; and decreasing the perception that parental literacy was a barrier for learning. Treatment effects on parental self-efficacy beliefs were again consistently similar for girls and boys.

6.5 Robustness Checks

On February 27, 2020, the first COVID-19 case appeared in Nigeria (Federal Ministry of Health 2020). One month later, the Nigerian government mandated the closure of all schools at the national level (UNESCO 2020). School closures not only affected our main outcomes of interest (i.e., proportion of out-of-school children and learning outcomes), but also affected data collection as the research team did not want to put at risk respondents and enumerators. As a result of COVID, data collection took place in two parts: March 5-21 and September 5-December 5, 2020, as shown in Figure 2 of Section 5. For the first data collection, we achieved 21% of data collection and the remaining 79% was done in the second stage.[16] As data collection followed the order of the implementation of the intervention, the first three weeks of March covered 36% of $T1$ participants (817 obs.), 24% of control (1,115 obs.) and no participant allocated to $T2$.

In this section, we explore to what extent our main conclusions on learning outcomes remain the same after applying various robustness checks. For out-of-school children, this analysis is not needed as we use a retrospective measure of school attendance based on March 2020. Thus, this section focuses on literacy and numeracy skills. We distinguish three types of samples: *Sample 1*) $T1$ and control participants who were interviewed in March 2020, *Sample 2*) $T1$, $T2$ and control participants interviewed in March 2020 and participants living in catchment areas where less than 10% of respondents reported that their child is not currently attending school due to school closures because of COVID-19 and *Sample 3*) T1, T2 and control participants interviewed in March and in November-December 2020 when most of the schools were opened. We present balance tests for all three samples in the Appendix Tables A22, A23 and A24. We do not observe specific patterns characterizing one sample compared to the others. On average, the different subsamples characterizing T, $T1$ and $T2$ under the three definitions show statistical balance between treatments and control group.

Table 7 shows treatment effects for *Sample 1*, the sample of participants that were not exposed to school closures. Similar to Table 3 in Section 6.1, we observe that learning outcomes were not significantly affected by the aspirational videos. However, the sign of

[16]For this second stage, the authors obtained an additional ethical approval on September 3, 2020 from the IRB Solutions where the authors submitted a detailed protocol to the IRB to minimize the spread of COVID during data collection.

the effects are moving in the expected direction – we observe a decrease in the proportion of zeros in letter recognition and small increases in literacy and numeracy skills. Because we only have 21% of the study sample, we are not surprised about the lack of statistical power for the estimated treatment effect.

Table 7: Treatment Effects for Target Child on Literacy and Numeracy Test Scores: Pre-Covid *Sample 1* March 2020

Variables	*Literacy Skills*		*Numeracy Skills*
	Prop. Zero	Z-score	Z-score
	(1)	(2)	(3)
Videos (T1)	-0.025	0.034	0.035
	(0.049)	(0.109)	(0.107)
Observations	1,812	1,812	1,812
R-squared	0.265	0.369	0.320
Covariates	YES	YES	YES
Enumerator FE	YES	YES	YES
Baseline Mean	0.956	0.000	0.000
Control Mean in F-up	0.685	-0.715	0.318

Note: Column 1 shows the results of linear probability models of the proportion of children scoring "zero" in letter recognition. Column 2-3 show z-scores of literacy and numeracy skills, respectively.Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 8 shows that *Sample 2* and *Sample 3* provide the same conclusion as our main outcome in Table 3. Overall, aspirational videos alone ($T1$) did not have a significant impact on learning outcomes. It was the combination of both interventions that had a large impact on literacy and numeracy skills for the target child. In addition, we also observe that the impact on learning outcomes is slightly higher for the sub-samples a smaller proportion of school closures. This suggests potential complementarities of school activities and home learning using the mobile apps.

Table 8: Treatment Effect for Target Child on Literacy and Numeracy Test Scores: Pre- and Post-Covid Samples

	Literacy Skills				*Numeracy Skills*	
	Prop. Zero		Z-score		Z-score	
	(1)	(2)	(3)	(4)	(5)	(6)
	Sample 2 Pre- and Post-Covid					
	March 2020 & Sept-Dec with less than 10% of school closures					
Any Treatment (T1 or T2)	-0.094*		0.274**		0.439**	
	(0.053)		(0.133)		(0.178)	
Videos (T1)		0.121*		-0.129		-0.328
		(0.071)		(0.127)		(0.207)
Videos + Mobile apps (T2)		-0.218***		0.507***		0.880***
		(0.048)		(0.160)		(0.191)
Observations	5,077	5,077	5,077	5,077	5,077	5,077
R-squared	0.200	0.243	0.296	0.302	0.285	0.312
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Baseline Mean	0.977	0.977	0.000	0.000	0.000	0.000
Control Mean in F-up	0.453	0.453	-0.684	-0.684	1.101	1.101
T1 vs T2 (p-value)		0.000		0.000		0.000
	Sample 3 Pre- and Post-Covid					
	March and Nov-Dec 2020 (most of the schools opened)					
Any Treatment (T1 or T2)	-0.075		0.105		0.249	
	(0.052)		(0.128)		(0.164)	
Videos (T1)		0.084		-0.161		-0.295
		(0.062)		(0.126)		(0.192)
Videos + Mobile apps (T2)		-0.228***		0.361**		0.773***
		(0.050)		(0.172)		(0.203)
Observations	4,863	4,863	4,863	4,863	4,863	4,863
R-squared	0.203	0.242	0.285	0.289	0.287	0.311
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Baseline Mean	0.974	0.974	0.000	0.000	0.000	0.000
Control Mean in F-up	0.479	0.479	-0.753	-0.753	1.013	1.013
T1 vs T2 (p-value)		0.000		0.002		0.000

Note: Columns 1 and 2 show the results of linear probability models of the proportion of children scoring "zero" in letter recognition. Column 3-4 show z-scores of literacy skills and Column 5-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

6.6 Effectiveness and Cost-Effectiveness

Our effectiveness and cost-effectiveness analyses compare the community impact of $T2$ (aspirational videos and m-Learning add-on) with other interventions aimed at improving learning outcomes in developing countries. We use the treatment effects of $T2$ as it was the intervention that had a significant impact on learning outcomes. Using the metric proposed by Angrist et al. (2020b), we calculated learning-adjusted years of schooling (LAYS). This metric combines educational access and quality; and it allows the comparison across countries using an absolute benchmark. To calculate the effectiveness and cost-effectiveness for this study, we took the following two steps.

First, we converted our aggregate measures of literacy and numeracy skills in LAYS. To facilitate the comparison of our effectiveness and cost-effectiveness results using the same list of educational interventions, this first effectiveness list already uses the 72 interventions that included cost data in Angrist et al. (2020b).[17]

Second, we use the cost of our combined intervention to calculate a cost-effectiveness measure that we adjust to country-specific price levels. For this step, we used nominal cost data which is subsequently transformed into 2020 purchasing power parity (PPP) adjusted prices, following JPAL Conducting Cost-effectiveness Analysis guidelines.[18] We then calculated the implied LAYS per US$100 (nominal) and PPP adjusted LAYS per US$100. Finally, we use the latter to compare the cost-effectiveness of our intervention with 72 other interventions, including edtech interventions.

Following the above steps, the treatment effects of this study for the target child (0.46σ and 0.63σ for literacy and numeracy, respectively) correspond to 0.57 and 0.79 LAYS, respectively. These metrics fall short of an artificial high-performance benchmark of 0.80, chosen by Angrist et al. (2020b) due to its stability and non-political properties. In terms of relative effectiveness, $T2$ ranks $5th$ relative to the 72 studies we consider for this comparison. Table 9 shows the twelve interventions with the largest impacts on LAYS. This list includes studies of two edtech interventions evaluated in India, computer-assisted learning by Banerjee et al. (2007) and vouchers for the Mindspark adaptive learning software by Muralidharan et al. (2019).[19]

Considering that $T2$ had similar effects on literacy and numeracy outcomes for older

[17] The study of Angrist et al. (2020b) uses data of 150 impact evaluations for its effectiveness estimates. It then calculates a cost-effectiveness measure for a subset of 72 interventions for which cost data were available. We compare our effectiveness and cost-effectiveness measures to the same list of 72 interventions.

[18] See more details about this methodology at `https://www.povertyactionlab.org/resource/conducting-cost-effectiveness-analysis-cea`.

[19] Our $T2$ and these two edtech interventions in India share characteristics that Angrist et al. (2020b) note are common among top performing educational interventions, including being multidimensional combined programs and instruction adjusted to the students' learning level.

siblings in treated communities (and many other spillovers), the measure of LAYS for $T2$ is likely to be a conservative estimate of the effectiveness of our combined intervention.

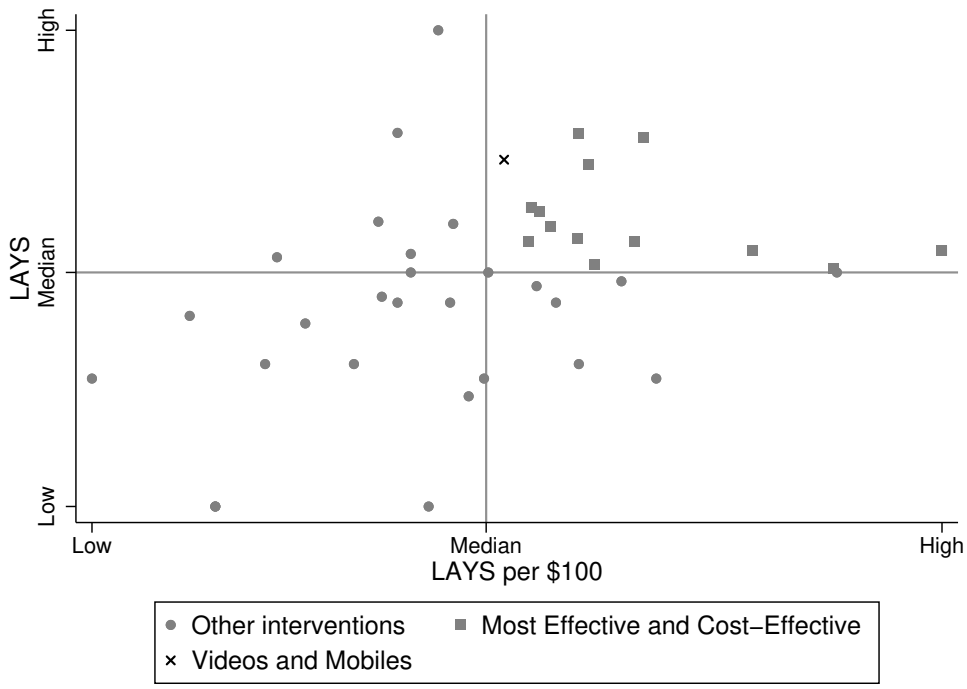
Table 9: Most effective educational interventions, out of 72 impact evaluations with cost data

Rank	Intervention	Country	LAYS
1	Combined intervention: improved pedagogy para-teachers and targeted instruction	The Gambia	4
2	Combined scholarships for girls, school materials and training for teachers and parents	Tanzania	1.1
3	Psycho-social and other support through school-community engagement including training of teacher mentors, parent support groups and local community authorities	Tanzania	1.09
4	"Let's Read" program, that provides textbooks, teacher coaching, and teacher training	Kenya	1.04
5	**Videos and Mobiles intervention (this study)**	**Nigeria**	**0.57-0.79**
6	Building village-based schools	Afghanistan	0.74
7	Computer-assisted learning (CAL)	India	0.43
8	Community-based preschools	Mozambique	0.41
9	Vouchers for Mindspark adaptive learning software	India	0.36
10	Preschool attendance	Uruguay	0.35
11	Merit scholarships for girls	Kenya	0.34
12	Ability grouping using extra teachers	Kenya	0.29

When we consider the cost of delivering the aspirational videos through community screenings and providing the preloaded smartphones and solar chargers (the cost of the videos and apps are excluded, both funded through multi-donor initiatives and publicly distributed), the cost per child in $T2$ was estimated at US$76.4. This nominal cost was converted into LAYS/$100 which corresponded to 0.75 and 1.02 for literacy and numeracy, respectively. Therefore, when considering both effectiveness and cost, $T2$ ranks in the top quarter of the 72 interventions we used to compare the treatment effects.

See figure 3 for a visual representation of our effectiveness and cost-effectiveness results. Our five-day intervention (2 days during the weekend, one day to explain the use of m-Learning apps and 2 days to monitor mobile usage) falls in the upper-right quadrant, greatly above the medians for effectiveness (LAYS) and cost-effectiveness (LAYS per US$100).

Figure 3: Cost-Effectiveness of Videos and Mobiles: LAYS vs LAYS per US $100



7 Conclusions

School enrollment and learning outcomes often lag far behind in settings where traditional social norms prevail. This is observed in northern Nigeria, where Shari'a law is formally institutionalized, and many families have adverse attitudes towards education. Low literacy rates and poor quality of formal education are also important barriers in low-resource settings. Considering the very high rates of learning poverty in Sub-Saharan Africa, being further worsened by the COVID-19 pandemic, it is important to test scalable and cost-effective home-learning initiatives (The World Bank 2022).

This paper presented experimental evidence of two innovations that targeted 6-9 year olds and their parents in northern Nigeria: aspirational videos delivered through community screenings and the provision of pre-loaded smartphones aimed at boosting the impacts of these videos. After an average of 12 months, we find that this five-day intervention had large and statistically significant effects on our main outcomes. Our intent-to-treat estimates show that the aspirational videos where very effective at reshaping parental educational attitudes and decreasing the proportion out-of-school children by 42 percent. Despite the study taking place in communities where local schools had recently received large supply-side investments, increased enrollment did not lead to improved

learning outcomes, highlighting the limitations of supply-side investments observed in other studies in developing countries (see Conn (2017) and Evans and Mendez Acosta (2021)).

On the other hand, our results suggest the high potential of investing in edtech home-learning approaches. Consistent with this line of thought, when aspirational videos were complemented with the m-Learning component in treatment communities, children's literacy and numeracy skills respectively improved by 0.46σ and 0.63σ in comparison to the control group. These large effects are lower-bound estimates for a home-learning intervention considering that only 40% of households attending the community screenings received a mobile phone with m-Learning apps.

The spillover effects of the m-Learning add-on underscore the fact that smartphones are often used by multiple household members in low-resource settings. Their provision improved literacy and numeracy skills in a similar magnitude for older, non-targeted siblings; reduced early parenthood of teenagers living in the treated households; and decreased early insertion into the labor market. When we compare treatment effects between participants who received (40%) and did not received a mobile phone (60%) living in the communities selected for both the videos and the m-Learning add-on, our results show that these effects were 75% as large as the effects observed on target children directly benefiting from the device.

Our data suggests that both components were complementary to each other, though the m-Learning add-on often played a more prominent role in our low-literate context. This booster arm greatly drove impacts across multiple outcomes, including improved parental self-perception about their capabilities to help their children to learn and perceptions about their children's own capabilities. The m-Learning add-on also had impact on reducing early entrance into the labor market. However, the aspirational videos alone were effective at decreasing out-of-school children and reducing parental desires for early marriage of their children. While both components were effective in increased setting time and space for study (parents reading to the target child, making time for school and learning activities, and helping their child develop communication skills by naming objectives) and in reshaping parental perceptions of prevalent social norms in their communities related to schooling and early marriage, most of these mechanisms were greatly driven by the edtech arm.

We generally find no evidence of heterogeneous effects by the child's gender, highlighting the potential of edtech to also effectively reach girls in conservative settings, where girls' seclusion or a strong bias towards boys' education may prevent girls from accessing formal schooling. Our heterogeneous analysis by gender shows that the in-

terventions worked for both girls and boys, and that the magnitude of treatment effects across gender were generally similar for the main outcomes (school enrollment, and literacy/numeracy skills). However, we observe gender differential impacts in terms of aspirations, as these greatly differ for boys and girls in our study setting. Parental aspirations were largely impacted by the aspirational videos alone ($T1$) for parents with daughters; on a smaller scale, aspirational videos with the m-Learning ($T2$) add-on changed parental aspirations for parents with sons (but not $T1$ alone). The aspirational videos alone also decreased aspirations for seeing daughters getting married at the age of 15 and 18 (we do not find effects for boys, as they marry much later than girls). These findings show the need for educational interventions to address the inter-linkages of schooling and marriage decisions as schooling may be perceived as a deterrent of social practices highly valued in traditional societies.

When evaluating the effectiveness of the combined intervention with 72 other educational interventions, we observe that this intervention ranked 5th in terms of learning gains. When considering costs, the combined intervention ranked in the upper quarter in terms of cost-effectiveness.

In conclusion, this study demonstrated that innovative approaches that use aspirational movies and engaging mobile apps can help remove major barriers to improve educational and development outcomes for households living in rural, low-literate communities that are governed by traditional social norms. Considering that the evidence for these traditional settings is limited, we hope our study motivates researchers and policymakers to continue testing home-learning strategies to improve investments in human capital in hard-to-reach populations. Such strategies are also needed for the broader population, where learning outcomes remained stagnated prior to the COVID-19 pandemic, and important losses are concentrating among marginalized populations and the poor.

References

Anderson, C., Hildreth, J. A. D., and Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological bulletin*, 141(3):574.

Angrist, N., Bergman, P., Brewster, C., and Matsheng, M. (2020a). Stemming learning loss during the pandemic: A rapid randomized trial of a low-tech intervention in botswana. Number 3663098 in SSRN.

Angrist, N., Evans, D. K., Filmer, D., Glennerster, R., Rogers, F. H., and Sabarwal, S. (2020b). How to improve education outcomes most efficiently? A comparison of 150 interventions using the new learning-adjusted years of schooling metric.

Antoninis, M. (2014). Tackling the largest global education challenge? Secular and religious education in Northern Nigeria. *World Development*, 59:82–92.

Araya, R., Arias Ortiz, E., Bottan, N., and Cristia, J. (2019). Does gamification in education work. *Experimental Evidence from Chile. Available online: https://publications.iadb.org/publications/english/document/Does_Gamification_in_Education_Work_Experimental_Evidence_from_Chile_en_en.pdf (accessedon23June2020)*.

Arias, E. (2019). How does media influence social norms? Experimental evidence on the role of common knowledge. *Political Science Research and Methods*, 7(3):561–578.

Azomahou, T. T. and Yitbarek, E. A. (2016). Intergenerational education mobility in Africa: has progress been inclusive? *World Bank Policy Research Working Paper*, 7843.

Baird, S., McIntosh, C., and Özler, B. (2019). When the money runs out: Do cash transfers have sustained effects on human capital accumulation? *Journal of Development Economics*, 140:169–185.

Bandura, A. (1971). Vicarious and self-reinforcement processes. *The nature of reinforcement*, 228278.

Bandura, A. (1997). *Self-Efficacy The Exercise of Control*. W.H. Freeman and Company, New York.

Bandura, A. (2003). Social cognitive theory for personal and social change by enabling media. In *Entertainment-education and social change*, pages 97–118. Routledge.

Banerjee, A. V., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., and Walton, M. (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India. Technical report, National Bureau of Economic Research.

Banerjee, A. V., Cole, S., Duflo, E., and Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, 122(3):1235–1264.

Banerjee, A. V. and Duflo, E. (2011). Why aren't children learning. Technical Report 1.

Banerjee, A. V., La Ferrara, E., and Orozco-Olvera, V. H. (2019). The entertaining way to behavioral change: Fighting HIV with MTV. Technical report, National Bureau of Economic Research.

Berg, G. and Zia, B. (2017). Harnessing emotional connections to improve financial decisions: Evaluating the impact of financial education in mainstream media. *Journal of the European Economic Association*, 15(5):1025–1055.

Berlinski, S., Galiani, S., and Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, 93(1-2):219–234.

Bertoni, E., Di Maio, M., Molini, V., and Nistico, R. (2019). Education is forbidden: The effect of the Boko Haram conflict on education in North-East Nigeria. *Journal of Development Economics*, 141:102249.

Blair, G., Littman, R., and Paluck, E. L. (2019). Motivating the adoption of new community-minded behaviors: An empirical test in Nigeria. *Science Advances*, 5(3):1–8.

Brunette, T., Piper, B., Jordan, R., King, S., and Nabacwa, R. (2019). The impact of mother tongue reading instruction in twelve Ugandan languages and the role of language complexity, socioeconomic factors, and program implementation. *Comparative Education Review*, 63(4):591–612.

Conn, K. M. (2017). Identifying effective education interventions in Sub-Saharan Africa: A meta-analysis of impact evaluations. *Review of Educational Research*, 87(5):863–898.

Duflo, E. (2004). The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia. *Journal of Development Economics*, 74(1):163–197.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5):1739–74.

Evans, D. K. and Mendez Acosta, A. (2021). Education in Africa: What are we learning? *Journal of African Economies*, 30(1):13–54.

Evans, D. K. and Popova, A. (2016). What really works to improve learning in developing countries? *World Bank Research Observer*, 31(2):242–270.

Federal Ministry of Health (2020). Health Minister: First Case of COVID-19 confirmed in Nigeria. `https://www.health.gov.ng/index.php?option=com_k2&view=item&id=613:health-minister-first-case-of-covid-19-confirmed-in-nigeria`. [Online; accessed 18-July-2022].

Ganimian, A. J. and Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, 86(3):719–755.

GEEAP (2020). Cost effective approaches to improve global learning: What does recent evidence tell us are "smart buys" for improving learning in low- and middle-income countries? Technical report, The World Bank, UK Foreign, Commonwealth & Development Office and Building Evidence in Education.

Glewwe, P., Kremer, M., and Moulin, S. (2009). Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, 1(1):112–135.

Grady, C., Iannantuoni, A., and Winters, M. S. (2021). Influencing the means but not the ends: The role of entertainment-education interventions in development. *World Development*, 138:105200.

Green, D. P., Wilke, A. M., and Cooper, J. (2020). Countering violence against women by encouraging disclosure: A mass media experiment in Rural Uganda. *Comparative Political Studies*, 53(14):2283–2320.

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., and Puranen, B. (2022). World Values Survey: Round seven—country-pooled datafile version 5.0. *JD Systems Institute & WVSA Secretariat: Madrid, Spain and Vienna, Austria*.

Human Right Watch (2014). "Political Shari'a"? Human rights and Islamic law in Northern Nigeria. Technical Report 9A.

Ito, H., Kasai, K., Nishiuchi, H., and Nakamuro, M. (2019). Does computer-aided instruction improve children's cognitive and non-cognitive skills? *Asian Development Review*, 38(1):98–118.

Jensen, R. and Oster, E. (2009). The Power of TV: Cable Television and Women's Status in India. *The Quarterly Journal of Economics*, 124(3):1057–1094.

La Ferrara, E., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Mbiti, I. M. (2016). The need for accountability in education in developing countries. *Journal of Economic Perspectives*, 30(3):109–32.

Moussa, W. and Omoeva, C. (2020). The long-term effects of universal primary education: Evidence from Ethiopia, Malawi, and Uganda. *Comparative Education Review*, 64(2):179–206.

Muralidharan, K., Singh, A., and Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, 109(4):1426–60.

NPC and ICF (2019). Nigeria Demographic and Health Survey 2018. Technical report, National Population Commission [Nigeria] and ICF.

Orozco-Olvera, V., Shen, F., and Cluver, L. (2019). The effectiveness of using entertainment education narratives to promote safer sexual behaviors of youth: A meta-analysis, 1985-2017. *PLOS ONE*, 14(2):1–14.

Popova, A., Evans, D., Breeding, M. E., and Arancibia, V. (2018). Teacher professional development around the world: The gap between evidence and practice. *World Bank Policy Research Working Paper*, (8572).

Razzu, G. and Wambile, A. (2021). Four decades of intergenerational educational mobility in Sub-Saharan Africa. *The Journal of Development Studies*, pages 1–20.

Rodriguez-Segura, D. (2021). Edtech in developing countries: A review of the evidence. *The World Bank Research Observer*, 37(2):171–203.

Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., Grau, V., Lagos, F., López, X., López, V., et al. (2003). Beyond Nintendo: design and assessment of educational video games for first and second grade students. *Computers & Education*, 40(1):71–94.

Schultz, T. P. (2004). School subsidies for the poor: Evaluating the Mexican Progresa poverty program. *Journal of Development Economics*, 74(1):199–250.

Suberu, R. T. (2009). Religion and institutions: Federalism and the management of conflicts over Sharia in Nigeria. *Journal of International Development*, 21(4):547–560.

The World Bank (2021). Primary school enrolment (% gross) @ONLINE. `https://data.worldbank.org/country/nigeria?view=chart`.

The World Bank (2022). The state of global learning poverty: 2022 update conference edition. Technical report, The World Bank, UNESCO, UNICEF, USAID, FCDO, Bill and Melinda Gates Foundation.

UNESCO (2019). New methodology shows that 258 million children, adolescents and youth are out of school. `https://unesdoc.unesco.org/ark:/48223/pf0000382577`. [Online; accessed 14-October-2021. Factsheet 56, UIS/2019/ED/FS/56].

UNESCO (2020). Dashboards on the Global Monitoring of School Closures Caused by the COVID-19 Pandemic. `https://covid19.uis.unesco.org/global-monitoring-school-closures-covid19/`. [Online; accessed 18-July-2022].

UNESCO (2022). New estimation confirms out-of-school population is growing in Sub-Saharan Africa. `https://unesdoc.unesco.org/ark:/48223/pf0000382577`. [Online; accessed 5-February-2023. Factsheet 62, Policy paper 48].

USAID (2019). Reading and Access Research Activity (RARA) Final Activity Report. Technical report.

Appendix

Table A1: Statistical Balance at Baseline: Mean differences between treatment groups and control

Variables	T vs C	SE	P-value	T1 vs C	SE	P-value	T2 vs C	SE	P-value
	Socio-Demographic Characteristics								
Female	-0.004	[0.01]	0.52	0.006	[0.01]	0.55	0.003	[0.01]	0.64
Age of target child	-0.010	[0.03]	0.69	-0.009	[0.03]	0.78	0.030	[0.03]	0.36
Age of father	-0.083	[0.38]	0.83	-0.137	[0.47]	0.77	0.296	[0.47]	0.53
Age of mother	-0.034	[0.28]	0.90	-0.087	[0.35]	0.80	0.153	[0.33]	0.64
Household size	0.022	[0.16]	0.89	-0.039	[0.20]	0.85	-0.006	[0.19]	0.98
Number of rooms	-0.029	[0.07]	0.69	-0.013	[0.09]	0.88	0.071	[0.08]	0.41
Always eat enough	-0.014	[0.02]	0.50	0.013	[0.02]	0.58	0.014	[0.03]	0.59
Not always eat enough	0.015	[0.02]	0.54	-0.018	[0.03]	0.56	-0.012	[0.03]	0.69
Frequently not eat enough	-0.002	[0.02]	0.91	0.005	[0.02]	0.80	-0.002	[0.02]	0.92
HH education: none	0.028	[0.03]	0.29	-0.015	[0.03]	0.64	-0.041	[0.03]	0.20
HH education: Islamiyya / other	-0.012	[0.01]	0.15	0.009	[0.01]	0.39	0.014	[0.01]	0.13
HH education: any primary	0.006	[0.02]	0.69	-0.002	[0.02]	0.93	-0.010	[0.02]	0.55
HH education: any secondary	-0.013	[0.01]	0.38	0.010	[0.02]	0.55	0.015	[0.02]	0.41
HH education: higher than secondary	-0.009	[0.01]	0.51	-0.002	[0.01]	0.88	0.021	[0.02]	0.30
Number of books	0.013	[0.03]	0.64	0.005	[0.03]	0.89	-0.030	[0.03]	0.33
More than 20 books	0.003	[0.03]	0.91	-0.025	[0.03]	0.46	0.018	[0.04]	0.61
Less than 20 books	-0.016	[0.02]	0.43	0.020	[0.02]	0.41	0.012	[0.02]	0.62
Language spoken: Hausa	-0.025	[0.03]	0.42	0.034	[0.03]	0.28	0.016	[0.04]	0.69
Toilet: flush to piped sewer	-0.008	[0.01]	0.12	0.009	[0.01]	0.26	0.007	[0.01]	0.22
Toilet: pit latrine (VIP)	-0.050**	[0.02]	0.04	0.034	[0.03]	0.23	0.065*	[0.03]	0.06
Toilet: pit latrine w/slab	-0.029	[0.04]	0.49	0.004	[0.05]	0.93	0.053	[0.05]	0.27
Toilet: pit latrine without slab/open pit	0.042	[0.05]	0.40	-0.050	[0.06]	0.41	-0.034	[0.06]	0.57
Toilet: no facility/bush/field	0.072*	[0.04]	0.08	-0.030	[0.05]	0.57	-0.113***	[0.04]	0.01
Toilet: other-ventilated, hanging and other	-0.027*	[0.02]	0.09	0.033	[0.03]	0.19	0.021	[0.02]	0.22
Water: piped	-0.001	[0.01]	0.95	0.012	[0.02]	0.54	-0.011	[0.01]	0.38
Water: public tap	0.019	[0.03]	0.51	-0.019	[0.04]	0.59	-0.019	[0.04]	0.59
Water: tubewell or borehole	0.014	[0.04]	0.74	-0.031	[0.05]	0.57	0.003	[0.05]	0.94
Water: protected dug well	-0.013	[0.01]	0.20	0.006	[0.01]	0.62	0.021	[0.01]	0.14
Water: unprotected dug well	-0.028	[0.04]	0.54	0.028	[0.06]	0.63	0.028	[0.05]	0.60
Water: protected/unprotected spring	0.000	[0.00]	0.90	0.001	[0.00]	0.57	-0.001	[0.00]	0.67
Water: surface water/river/canal/irrigation	0.014	[0.02]	0.37	-0.004	[0.02]	0.82	-0.023	[0.01]	0.12
Water: other source	-0.005	[0.00]	0.35	0.008	[0.01]	0.40	0.002	[0.00]	0.62
	Main Outcomes: School Attendance, Uwezo, EGRA and EGMA of target child								
Child currently attends formal school	0.005	[0.03]	0.85	-0.004	[0.03]	0.91	-0.006	[0.03]	0.84
Letter recognition (Uwezo)	0.028	[0.08]	0.74	-0.107	[0.09]	0.23	0.050	[0.11]	0.66
Word reading (Uwezo)	0.206	[0.26]	0.44	-0.174	[0.31]	0.58	-0.232	[0.31]	0.46
Sentence reading (Uwezo)	0.568	[0.69]	0.42	-1.096	[0.86]	0.21	-0.146	[0.84]	0.86
Reading comprehension (Uwezo)	0.262	[0.40]	0.52	-0.462	[0.80]	0.57	-0.128	[0.38]	0.74
Letter recognition (EGRA)	0.062	[0.11]	0.58	-0.068	[0.12]	0.58	-0.056	[0.14]	0.70
Nonword reading (EGRA)	1.796	[2.06]	0.39	-2.317	[1.84]	0.22	-1.432	[2.68]	0.60
Oral passage reading (EGRA)	0.030	[0.06]	0.60	0.005	[0.05]	0.92	-0.064	[0.08]	0.42
Reading comprehension (EGRA)	0.002	[0.01]	0.86	-0.012	[0.01]	0.17	0.008	[0.02]	0.73
Listening comprehension	-0.059	[0.13]	0.64	0.021	[0.15]	0.89	0.096	[0.16]	0.54
Letter dictation (EGRA)	0.006	[0.08]	0.95	-0.019	[0.10]	0.85	0.007	[0.11]	0.95
Word dictation: correct word (EGRA)	0.006	[0.02]	0.69	-0.004	[0.02]	0.82	-0.009	[0.02]	0.69
Word dictation: all letters correct (EGRA)	0.007	[0.02]	0.64	-0.003	[0.02]	0.83	-0.011	[0.02]	0.62
Number recognition (EGMA)	-0.017	[0.08]	0.83	0.006	[0.10]	0.96	0.029	[0.10]	0.78
Number discrimination (EGMA)	0.014	[0.22]	0.95	-0.115	[0.27]	0.67	0.085	[0.24]	0.73
Missing number (EGMA)	-0.007	[0.04]	0.85	-0.001	[0.05]	0.99	0.015	[0.05]	0.75
Addition (EGMA)	-0.065	[0.06]	0.29	0.082	[0.07]	0.27	0.049	[0.08]	0.54
Substraction (EGMA)	-0.037	[0.05]	0.43	0.058	[0.05]	0.29	0.016	[0.06]	0.78

Note: HH refers to household head. OLS regressions with clustered standard errors at the school catchment area level. Dependent variable corresponds to each variable shown in this table at baseline on a dummy identifying T, T1 and T2, respectively. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A2: Literacy: Treatment Effect on Target Child by Specific Competence Assessed

	Letter Recognition	Passage Reading	Reading Compreh.	Listening Compreh.	Letter Dictation	Word Dictation	Word Dictation (all letters)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Videos (T1)	-0.087	0.012	-0.004	-0.041	-0.029	-0.101	-0.059
	(0.141)	(0.079)	(0.027)	(0.078)	(0.070)	(0.068)	(0.070)
Videos + Mobile apps (T2)	0.423**	0.183*	0.032	0.232***	0.408***	0.164**	0.111
	(0.183)	(0.104)	(0.035)	(0.063)	(0.078)	(0.082)	(0.087)
Observations	8,928	8,928	8,928	8,928	8,928	8,928	8,928
R-squared	0.160	0.069	0.032	0.279	0.276	0.136	0.130
Control Baseline Mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean in F-up	1.303	-5.421	0.035	-0.125	0.206	0.265	0.331
T1 vs T2 (p-value)	0.003	0.090	0.261	0.000	0.000	0.001	0.058

Note: OLS regressions with clustered standard errors at the school catchment area level. Test scores are normalized using the mean and standard deviation for the control group at baseline per specific competence. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A3: Numeracy: Treatment Effect on Target Child by Specific Competence Assessed

	Number Recognition	Number Discrimination	Missing Number	Addition	Subtraction
	(1)	(2)	(3)	(4)	(5)
Videos (T1)	-0.113	-0.131	-0.021	0.016	-0.059
	(0.160)	(0.099)	(0.092)	(0.077)	(0.063)
Videos + Mobile apps (T2)	0.667***	0.431***	0.474***	0.324***	0.325***
	(0.165)	(0.074)	(0.108)	(0.102)	(0.091)
Observations	8,859	8,928	8,928	8,928	8,928
R-squared	0.248	0.268	0.232	0.191	0.176
Control Baseline Mean	0.000	0.000	0.000	0.000	0.000
Control Mean in F-up	1.364	0.759	0.437	0.466	0.415
T1 vs T2 (p-value)	0.000	0.000	0.000	0.004	0.000

Note: OLS regressions with clustered standard errors at the school catchment area level. Test scores are normalized using the mean and standard deviation for the control group at baseline per specific competence. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A4: Treatment Effect on Parental Aspirations for Target Child: School, Work, Marriage and Apprenticeship

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	School		Marriage		Work		Apprentice	
				Aspirations at age 15				
Any Treatment (T1 or T2)	0.037**		-0.032***		-0.000		-0.005	
	(0.014)		(0.012)		(0.003)		(0.005)	
Videos (T1)		0.020		-0.025*		0.005		0.000
		(0.017)		(0.014)		(0.004)		(0.006)
Videos + Mobile apps (T2)		0.055***		-0.039***		-0.006**		-0.012**
		(0.015)		(0.013)		(0.002)		(0.005)
Observations	9,200	9,200	9,200	9,200	9,200	9,200	9,200	9,200
R-squared	0.174	0.175	0.207	0.207	0.027	0.029	0.028	0.029
Control Mean in F-up	0.826	0.826	0.139	0.139	0.011	0.011	0.023	0.023
T1 vs T2 (p-value)		0.028		0.303		0.003		0.017
				Aspirations at age 18				
Any Treatment (T1 or T2)	0.052***		-0.037***		-0.004		-0.013*	
	(0.017)		(0.013)		(0.004)		(0.007)	
Videos (T1)		0.030		-0.030*		0.003		-0.006
		(0.020)		(0.016)		(0.004)		(0.008)
Videos + Mobile apps (T2)		0.076***		-0.046***		-0.012***		-0.020***
		(0.021)		(0.017)		(0.003)		(0.007)
Observations	9,200	9,200	9,200	9,200	9,200	9,200	9,200	9,200
R-squared	0.314	0.315	0.383	0.383	0.024	0.026	0.062	0.062
Control Mean in F-up	0.649	0.649	0.281	0.281	0.015	0.015	0.051	0.051
T1 vs T2 (p-value)		0.046		0.398		0.000		0.065

Note: Linear probability models with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A5: Treatment Effect on Parental Aspirations for Target Child: Age to get married and to have first child

	(1)	(2)	(3)	(4)
	Aspired age to see their children			
Variables	Getting married		Having first child	
Any Treatment (T1 or T2)	0.444***		0.387***	
	(0.151)		(0.141)	
Videos (T1)		0.323**		0.330**
		(0.162)		(0.151)
Videos + Mobile apps (T2)		0.577***		0.450**
		(0.193)		(0.178)
Observations	9,167	9,167	9,167	9,167
R-squared	0.665	0.665	0.660	0.660
Control Mean in F-up	22.821	22.821	24.677	24.677
T1 = T2 (p-value)		0.187		0.481

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A6: Treatment Effect on Parental Expectations for Target Child: School, Work and Marriage

	Attending School by Jan/Feb 2021		Finishing						Work by Jan/Feb 2021		Getting Married by Jan/Feb 2021	
			Primary		Lower Secondary		High School					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Any Treatment (T1 or T2)	2.990***		2.333***		3.757***		4.830***		-0.644		-0.118	
	(0.754)		(0.630)		(1.090)		(1.572)		(1.273)		(0.123)	
Videos (T1)		2.595***		2.108***		3.604***		6.794***		1.505		0.048
		(0.835)		(0.724)		(1.182)		(1.622)		(1.525)		(0.155)
Videos + Mobile apps (T2)		3.420***		2.578***		3.924***		2.680		-2.983**		-0.298**
		(0.872)		(0.708)		(1.326)		(2.208)		(1.355)		(0.137)
Observations	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201
R-squared	0.127	0.127	0.107	0.107	0.162	0.162	0.313	0.315	0.262	0.265	0.010	0.010
Control Mean in F-up	92.217	92.217	94.248	94.248	86.509	86.509	59.663	59.663	14.531	14.531	0.436	0.436
T1 vs T2 (p-value)		0.309		0.492		0.797		0.063		0.004		0.039

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A7: Treatment Effect on Parental Self-Efficacy Beliefs: Capabilities to help improve or learn the target child

	Help learn/improve						Improve learning		Parental education/literacy an obstacle for			
	Literacy		Numeracy		Communication		By setting time & space		Literacy		Numeracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Any Treatment (T1 or T2)	1.945*		1.663		0.808		1.045		-0.011		-0.013	
	(1.073)		(1.135)		(1.034)		(1.060)		(0.013)		(0.013)	
Videos (T1)		-0.157		-0.818		-2.036		-1.738		0.012		0.011
		(1.340)		(1.413)		(1.244)		(1.277)		(0.017)		(0.017)
Videos + Mobile apps (T2)		4.233***		4.363***		3.904***		4.077***		-0.036***		-0.038***
		(0.961)		(1.013)		(0.947)		(0.978)		(0.012)		(0.012)
Observations	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201
R-squared	0.190	0.197	0.204	0.212	0.182	0.195	0.191	0.202	0.255	0.257	0.263	0.265
Control Mean in F-up	92.004	92.004	91.335	91.335	92.529	92.529	91.671	91.671	0.129	0.129	0.126	0.126
T1 vs T2 (p-value)		0.000		0.000		0.000		0.000		0.004		0.004

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A8: Treatment Effect on Parental beliefs about the capabilities of target children

	To Learn Skills in				Rate of Child's Ability in Comparison			
	Literacy		Numeracy		Literacy		Numeracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any Treatment (T1 or T2)	3.040**		3.295**		2.342**		2.296*	
	(1.297)		(1.323)		(1.170)		(1.177)	
Videos (T1)		2.467		2.651		1.294		1.409
		(1.582)		(1.632)		(1.407)		(1.441)
Videos + Mobile apps (T2)		3.665**		3.995***		3.483**		3.262**
		(1.509)		(1.510)		(1.341)		(1.341)
Observations	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201
R-squared	0.317	0.317	0.337	0.338	0.297	0.297	0.308	0.308
Control Mean in F-up	51.445	51.445	49.687	49.687	52.059	52.059	51.184	51.184
T1 vs T2 (p-value)		0.488		0.442		0.148		0.230

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A9: Time Allocation: Activities with/for target child

Variables	(1) Read to child	(2) Read to child	(3) Helped learn	(4) Helped learn	(5) Helped with homework or school work	(6) Helped with homework or school work	(7) Made time for school work/learning	(8) Made time for school work/learning	(9) Taught names of objects	(10) Taught names of objects	(11) Played an instrument for/with child	(12) Played an instrument for/with child	(13) Brought child to school	(14) Brought child to school
Any treatment (T1 or T2)	0.036***		0.029**		0.015		0.031**		0.030***		0.004		0.010	
	(0.013)		(0.014)		(0.009)		(0.013)		(0.011)		(0.005)		(0.008)	
Videos (T1)		0.024*		0.020		0.010		0.010		0.010		0.006		0.002
		(0.014)		(0.017)		(0.012)		(0.014)		(0.011)		(0.006)		(0.010)
Videos + Mobile apps (T2)		0.049***		0.039**		0.020*		0.053***		0.051***		0.003		0.020**
		(0.017)		(0.018)		(0.011)		(0.017)		(0.015)		(0.006)		(0.010)
Observations	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201
R-squared	0.079	0.080	0.050	0.050	0.031	0.032	0.045	0.047	0.125	0.127	0.034	0.034	0.049	0.050
Control Mean in F-up	0.130	0.130	0.106	0.106	0.036	0.036	0.070	0.070	0.091	0.091	0.025	0.025	0.038	0.038
T1 vs T2 (p-value)		0.179		0.375		0.476		0.022		0.008		0.625		0.099

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A10: Time Allocation: Activities done by the target child

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	Doing domestic chores		Reading / Studying at home		Attending islamiyya		Attending formal school		Working outside home to earn money	
Any treatment (T1 or T2)	-0.001		0.007		-0.001		0.220		-0.042	
	(0.079)		(0.030)		(0.064)		(0.323)		(0.029)	
Videos (T1)		0.051		0.031		-0.009		-0.330		0.017
		(0.096)		(0.038)		(0.089)		(0.415)		(0.032)
Videos + Mobile apps (T2)		-0.057		-0.020		0.007		0.820**		-0.105***
		(0.094)		(0.034)		(0.061)		(0.355)		(0.031)
Observations	9,071	9,071	9,071	9,071	9,127	9,127	9,172	9,172	9,071	9,071
R-squared	0.130	0.131	0.174	0.175	0.092	0.092	0.086	0.108	0.089	0.093
Baseline Mean	1.500	1.500	0.272	0.272	2.046	2.046	3.697	3.697	0.140	0.140
Control Mean in F-up	1.878	1.878	0.340	0.340	2.039	2.039	2.467	2.467	0.273	0.273
T1 vs T2 (p-value)		0.308		0.194		0.852		0.013		0.000

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A11: Social Norms about school and early entry to the labor market – Main Respondent and Village

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	School				Early entry to the labor market			
	Girls		Boys		Girls		Boys	
Any Treatment (T1 or T2)	2.963**		2.670**		-0.908		-1.041	
	(1.228)		(1.334)		(0.825)		(0.902)	
Videos (T1)		3.387**		2.914**		-1.465		-1.952**
		(1.350)		(1.400)		(1.025)		(0.948)
Videos + Mobile apps (T2)		2.503		2.404		-0.302		-0.050
		(1.647)		(1.779)		(1.004)		(1.255)
Observations	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201
R-squared	0.284	0.284	0.336	0.336	0.156	0.156	0.091	0.092
Baseline Mean	16.315	16.315	33.575	33.575	39.729	39.729	38.012	38.012
Control Mean in F-up	29.016	29.016	26.806	26.806	36.805	36.805	35.170	35.170
T1 vs T2 (p-value)		0.609		0.771		0.323		0.156

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A12: Social Norms about early marriage

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Main respondent & Village				Main respondent		Village	
	Girls		Boys		Girls			
Any Treatment (T1 or T2)	-2.389		0.109		-3.936***		-3.765**	
	(1.568)		(0.173)		(1.490)		(1.642)	
Videos (T1)		-0.516		0.396		-1.137		-0.476
		(1.755)		(0.243)		(1.661)		(1.811)
Videos + Mobile apps (T2)		-4.428**		-0.204		-6.981***		-7.346***
		(1.854)		(0.149)		(1.708)		(1.971)
Observations	9,201	9,201	9,201	9,201	9,201	9,201	9,201	9,201
R-squared	0.234	0.236	0.022	0.023	0.144	0.148	0.249	0.254
Baseline Mean	19.807	19.807	9.452	9.452	13.725	13.725	18.102	18.102
Control Mean in F-up	17.929	17.929	3.637	3.637	14.031	14.031	17.355	17.355
T1 vs T2 (p-value)		0.039		0.010		0.001		0.001

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A13: Spillovers on Siblings: Heterogeneous effects on Literacy and Numeracy Skills by Age

	Literacy Skills				*Numeracy Skills*	
	Prop. Zero		Z-score		Z-score	
	(1)	(2)	(3)	(4)	(5)	(6)
Age 13-18 (=1)	-0.204***	-0.205***	1.496***	1.498***	1.203***	1.205***
	(0.023)	(0.023)	(0.248)	(0.248)	(0.141)	(0.141)
Any Treatment (T1 or T2)	-0.062		0.095		0.121	
	(0.040)		(0.106)		(0.118)	
Any Treatment * Age 13-18 (=1)	0.102***		-0.275		-0.236	
	(0.036)		(0.345)		(0.188)	
Videos (T1)		0.017		-0.157		-0.209*
		(0.047)		(0.097)		(0.121)
Videos (T1) * Age 13-18 (=1)		0.113**		-0.014		-0.220
		(0.050)		(0.441)		(0.244)
Videos + Mobile apps (T2)		-0.151***		0.382***		0.497***
		(0.040)		(0.135)		(0.135)
Videos + M. apps (T2) * Age 13-18 (=1)		0.124***		-0.547		-0.372*
		(0.041)		(0.402)		(0.208)
Observations	5,268	5,268	5,268	5,268	5,268	5,268
R-squared	0.128	0.140	0.234	0.239	0.260	0.275
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Baseline Mean	0.942	0.942	0.000	0.000	0.000	0.000
Control Mean in F-up	0.476	0.476	0.932	0.932	1.127	1.127
T1 vs T2 (p-value)		0.000		0.000		0.000
T1 vs T1 13-18yo (p-value)		0.217		0.741		0.965
T2 vs T2 13-18yo (p-value)		0.000		0.033		0.001

Note: Siblings' age is defined according to their age at baseline. Column 1 and 2 show the results of linear probability models of the proportion of children scoring "zero" in letter recognition. Column 3-4 show z-scores of literacy skills and Column 5-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A14: Spillovers on Siblings: Heterogeneous effects on School Enrollment and Work by Age

	Attending School				*Work*			
	Currently		Between Jan-Mar2020		Past month		Between Jan-Mar2020	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age 13-18 (=1)	-0.114***	-0.113***	-0.198***	-0.130***	0.079***	0.074***	0.086***	0.081***
	(0.037)	(0.037)	(0.031)	(0.025)	(0.022)	(0.022)	(0.024)	(0.024)
Any Treatment (T1 or T2)	0.049		0.033		-0.002		0.001	
	(0.062)		(0.022)		(0.016)		(0.016)	
Any Treatment * Age 13-18 (=1)	0.011		-0.002		0.013		0.022	
	(0.050)		(0.043)		(0.034)		(0.036)	
Videos (T1)		-0.109		0.040**		0.010		0.012
		(0.078)		(0.017)		(0.022)		(0.022)
Videos (T1) * Age 13-18 (=1)		0.011		-0.116**		0.110**		0.131***
		(0.061)		(0.046)		(0.046)		(0.049)
Videos + Mobile apps (T2)		0.213***		0.022		-0.011		-0.009
		(0.057)		(0.017)		(0.017)		(0.016)
Videos + M. apps (T2) * Age 13-18 (=1)		-0.057		-0.018		-0.035		-0.034
		(0.050)		(0.040)		(0.033)		(0.036)
Observations	4,801	4,801	4,803	4,802	4,800	4,800	4,801	4,801
R-squared	0.143	0.187	0.060	0.262	0.123	0.130	0.121	0.128
Covariates	YES	YES	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES	YES	YES
Baseline Mean	0.814	0.814	0.814	0.814	0.024	0.024	0.024	0.024
Control Mean in F-up	0.492	0.492	0.813	0.813	0.135	0.135	0.135	0.135

Note: Siblings' age is defined according to their age at baseline. Column 1-8 and 2 show the results of linear probability models of attending school and of working in the past. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A15: Spillovers on Siblings: Literacy and Numeracy Skills in $T2$ communities

	Literacy Skills				*Numeracy Skills*	
	Prop. Zero		Z-score		Z-score	
	(1)	(2)	(3)	(4)	(5)	(6)
Any Treatment (T1 or T2)	-0.050		0.064		0.092	
	(0.038)		(0.116)		(0.122)	
Videos (T1)		0.032		-0.190*		-0.257**
		(0.046)		(0.115)		(0.123)
Videos + Mobile (T2)		-0.139***		0.355**		0.483***
		(0.042)		(0.174)		(0.154)
Videos + No Mobile (T2)		-0.138***		0.330**		0.462***
		(0.040)		(0.147)		(0.145)
Observations	5,268	5,268	5,268	5,268	5,268	5,268
R-squared	0.117	0.130	0.198	0.204	0.227	0.243
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Baseline Mean	0.942		0.000		0.000	
Control Mean in F-up	0.476	0.476	0.932	0.932	1.127	1.127
T2 mob. vs T2 no mob. (p-value)		0.958		0.849		0.833

Note: Column 1 and 2 show the results of linear probability models of the proportion of children scoring "zero" in letter recognition. Column 3-4 show z-scores of literacy skills and Column 5-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A16: Heterogeneous effects for Target Child: Out-of-School Children (Formal School)

	(1)	(2)
Female	0.014	0.014
	(0.011)	(0.011)
Any Treatment (T1 or T2)	-0.056***	
	(0.014)	
Any Treatment (T1 or T2) * Female	-0.015	
	(0.014)	
Videos (T1)		-0.055***
		(0.015)
Videos (T1) * Female		-0.015
		(0.015)
Videos + Mobile apps (T2)		-0.057***
		(0.016)
Videos + Mobile apps (T2) * Female		-0.015
		(0.016)
Observations	9,201	9,201
R-squared	0.239	0.239
Control Mean in Baseline (Girls)	0.170	0.170
Control Mean in Baseline (Boys)	0.165	0.165
Control Mean in F-up (Girls)	0.157	0.157
Control Mean in F-up (Boys)	0.139	0.139

Note: Linear probability models with clustered standard errors at the school catchment area level. The covariates considered in these regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A17: Heterogeneous Effects for Target Child by Gender: Literacy and Numeracy Test Scores

	Literacy Skills				*Numeracy Skills*	
	Prop. Zero		Z-score		Z-score	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.042***	0.042***	-0.077***	-0.076***	-0.089***	-0.089***
	(0.013)	(0.013)	(0.025)	(0.025)	(0.029)	(0.029)
Any Treatment (T1 or T2)	-0.073**		0.078		0.119*	
	(0.036)		(0.049)		(0.066)	
Any Treatment (T1 or T2) * Female	-0.008		-0.008		0.016	
	(0.020)		(0.040)		(0.043)	
Videos (T1)		0.023		-0.050		-0.084
		(0.044)		(0.050)		(0.076)
Videos (T1) * Female		0.003		0.043		0.044
		(0.027)		(0.044)		(0.050)
Videos + Mobile apps (T2)		-0.175***		0.211***		0.332***
		(0.034)		(0.060)		(0.074)
Videos + Mobile apps (T2) * Female		-0.021		-0.054		-0.007
		(0.023)		(0.051)		(0.054)
Observations	8,928	8,928	8,928	8,928	8,928	8,928
R-squared	0.163	0.184	0.248	0.253	0.264	0.280
Control Mean in Baseline (Girls)	0.982	0.982	-0.060	-0.060	-0.046	-0.046
Control Mean in Baseline (Boys)	0.963	0.963	0.062	0.062	0.048	0.048
Control Mean in F-up (Girls)	0.462	0.462	-0.067	-0.067	-0.061	-0.061
Control Mean in F-up (Boys)	0.412	0.412	0.070	0.070	0.064	0.064

Note: Column 1 and 2 show the results of linear probability models of the proportion of children scoring ″zero″ in letter recognition. Column 3-4 show z-scores of literacy skills and Column 5-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A18: Heterogeneous Effects by Gender and Age for Target Child: Literacy and Numeracy Test Scores

	Prop. Zero	Literacy Z-score	Numeracy Z-score	Prop. Zero	Literacy Z-score	Numeracy Z-score
	(1)	(2)	(3)	(4)	(5)	(6)
		Girls			Boys	
Age 7 at baseline	-0.098***	0.282***	0.348***	-0.072**	0.047	0.199**
	(0.024)	(0.092)	(0.085)	(0.028)	(0.126)	(0.092)
Age 8 at baseline	-0.148***	0.482***	0.478***	-0.103***	0.437***	0.731***
	(0.024)	(0.124)	(0.104)	(0.032)	(0.163)	(0.118)
Age 9 at baseline	-0.203***	0.863***	0.814***	-0.176***	0.742***	1.070***
	(0.032)	(0.177)	(0.121)	(0.040)	(0.195)	(0.154)
Videos (T1)	-0.042	0.154	0.057	0.062	-0.235*	-0.085
	(0.052)	(0.112)	(0.134)	(0.059)	(0.140)	(0.160)
Videos + Mobile apps (T2)	-0.247***	0.562***	0.728***	-0.220***	0.449**	0.725***
	(0.051)	(0.148)	(0.155)	(0.046)	(0.182)	(0.160)
Videos (T1)*Age 7 at baseline	0.122**	-0.218	-0.265*	-0.014	-0.003	-0.011
	(0.047)	(0.157)	(0.147)	(0.050)	(0.168)	(0.136)
Videos (T2)*Age 7 at baseline	0.069*	-0.342**	-0.364***	0.043	-0.011	-0.062
	(0.040)	(0.171)	(0.133)	(0.046)	(0.223)	(0.161)
Videos (T1)*Age 8 at baseline	0.088*	-0.349*	-0.198	-0.093**	0.129	-0.245
	(0.052)	(0.184)	(0.173)	(0.046)	(0.228)	(0.151)
Videos (T2)*Age 8 at baseline	0.035	-0.015	0.058	0.043	0.128	-0.166
	(0.043)	(0.200)	(0.185)	(0.046)	(0.277)	(0.240)
Videos (T1)*Age 9 at baseline	0.054	-0.094	-0.064	-0.064	0.340	-0.179
	(0.065)	(0.386)	(0.237)	(0.055)	(0.298)	(0.224)
Videos (T2)*Age 9 at baseline	0.099*	-0.168	0.110	0.042	0.555*	0.141
	(0.052)	(0.254)	(0.242)	(0.051)	(0.305)	(0.241)
Observations	4,590	4,590	4,590	4,338	4,338	4,338
R-squared	0.185	0.210	0.254	0.173	0.279	0.280
Covariates	YES	YES	YES	YES	YES	YES
Enumerator FE	YES	YES	YES	YES	YES	YES
Baseline Mean	0.982	-0.063	-0.049	0.970	0.066	0.052
Control Mean in F-up	0.462	-0.881	0.947	0.412	-0.537	1.186
T1 vs T2 (p-value)	0.000	0.013	0.000	0.000	0.000	0.000

Note: Column 1 and 4 show the results of linear probability models of the proportion of children scoring ″zero″ in letter recognition. Column 2 and 5 show z-scores of literacy skills and Column 3-6 z-scores of numeracy skills. Test scores are normalized using the mean and standard deviation for the control group at baseline. All regressions have clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A19: Treatment Effect on Parental Aspirations for Target Child by Gender

Variables	(1) School	(2) School	(3) Marriage	(4) Marriage	(5) Work	(6) Work	(7) Apprentice	(8) Apprentice
	Daughters							
	Aspirations at age 15							
Any Treatment (T1 or T2)	0.060***		-0.063***		-0.000		-0.000	
	(0.023)		(0.022)		(0.004)		(0.004)	
Videos (T1)		0.044		-0.054**		0.004		0.005
		(0.028)		(0.026)		(0.005)		(0.005)
Videos + Mobile apps (T2)		0.078***		-0.073***		-0.005*		-0.006
		(0.025)		(0.024)		(0.003)		(0.005)
Observations	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713
R-squared	0.130	0.130	0.124	0.124	0.045	0.046	0.026	0.027
Control Mean in F-up	0.702	0.702	0.271	0.271	0.010	0.010	0.016	0.016
T1 vs T2 (p-value)		0.192		0.461		0.045		0.051
	Aspirations at age 18							
Any Treatment (T1 or T2)	0.078***		-0.073***		-0.005*		-0.004	
	(0.025)		(0.025)		(0.003)		(0.006)	
Videos (T1)		0.063**		-0.062**		-0.003		-0.003
		(0.029)		(0.029)		(0.004)		(0.009)
Videos + Mobile apps (T2)		0.094***		-0.085***		-0.008***		-0.004
		(0.033)		(0.032)		(0.003)		(0.006)
Observations	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713
R-squared	0.146	0.146	0.125	0.125	0.028	0.028	0.051	0.051
Control Mean in F-up	0.411	0.411	0.547	0.547	0.010	0.010	0.029	0.029
T1 vs T2 (p-value)		0.393		0.529		0.058		0.970
	Sons							
	Aspirations at age 15							
Any Treatment (T1 or T2)	0.011		0.001		-0.001		-0.010	
	(0.009)		(0.001)		(0.004)		(0.006)	
Videos (T1)		-0.004		0.002		0.005		-0.003
		(0.011)		(0.002)		(0.005)		(0.008)
Videos + Mobile apps (T2)		0.027***		-0.000		-0.008**		-0.017***
		(0.008)		(0.001)		(0.004)		(0.006)
Observations	4,487	4,487	4,487	4,487	4,487	4,487	4,487	4,487
R-squared	0.059	0.062	0.021	0.021	0.036	0.038	0.044	0.045
Control Mean in F-up	0.955	0.955	0.001	0.001	0.011	0.011	0.031	0.031
T1 vs T2 (p-value)		0.002		0.248		0.006		0.032
	Aspirations at age 18							
Any Treatment (T1 or T2)	0.022*		0.002		-0.002		-0.022**	
	(0.013)		(0.002)		(0.006)		(0.010)	
Videos (T1)		-0.007		0.004		0.010		-0.008
		(0.015)		(0.003)		(0.008)		(0.011)
Videos + Mobile apps (T2)		0.054***		-0.000		-0.016***		-0.037***
		(0.013)		(0.002)		(0.005)		(0.011)
Observations	4,487	4,487	4,487	4,487	4,487	4,487	4,487	4,487
R-squared	0.095	0.101	0.023	0.024	0.038	0.043	0.088	0.090
Control Mean in F-up	0.898	0.898	0.004	0.004	0.020	0.020	0.074	0.074
T1 vs T2 (p-value)		0.000		0.104		0.000		0.012

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A20: Treatment Effect on Parental Expectations for Target Child by Gender

	Attending School by Jan/Feb 2021		Finishing Primary		Finishing Lower Secondary		Finishing High School		Work by Jan/Feb 2021		Getting Married by Jan/Feb 2021	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Daughters						
Any Treatment (T1 or T2)	3.064***		2.326***		4.976***		6.620***		-0.441		-0.099	
	(0.838)		(0.743)		(1.623)		(2.093)		(1.589)		(0.199)	
Videos (T1)		2.518***		1.919**		4.857***		10.737***		2.315		0.177
		(0.915)		(0.889)		(1.709)		(2.249)		(1.741)		(0.248)
Videos + Mobile apps (T2)		3.655***		2.769***		5.104**		2.144		-3.431*		-0.399*
		(1.010)		(0.847)		(2.040)		(2.923)		(1.832)		(0.202)
Observations	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713
R-squared	0.140	0.140	0.120	0.120	0.153	0.153	0.187	0.192	0.300	0.303	0.021	0.022
Control Mean in F-up	92.111	92.111	93.669	93.669	80.579	80.579	42.054	42.054	16.783	16.783	0.470	0.470
T1 vs T2 (p-value)		0.238		0.347		0.896		0.006		0.003		0.014
						Sons						
Any Treatment (T1 or T2)	2.789***		2.191***		2.341***		2.831*		-0.812		-0.137	
	(0.883)		(0.705)		(0.791)		(1.515)		(1.189)		(0.160)	
Videos (T1)		2.556**		2.136***		2.166**		2.584*		0.618		-0.096
		(0.977)		(0.761)		(0.864)		(1.545)		(1.574)		(0.194)
Videos + Mobile apps (T2)		3.043***		2.251***		2.533**		3.103		-2.370*		-0.182
		(1.020)		(0.822)		(0.971)		(1.954)		(1.221)		(0.205)
Observations	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488
R-squared	0.138	0.138	0.112	0.112	0.121	0.121	0.187	0.187	0.226	0.228	0.011	0.011
Control Mean in F-up	92.327	92.327	94.851	94.851	92.675	92.675	77.973	77.973	12.190	12.190	0.401	0.401
T1 vs T2 (p-value)		0.604		0.873		0.695		0.770		0.062		0.719

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of child, gender of child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A21: Treatment Effect on Parental Self-Efficacy Beliefs by Gender

	Help learn/improve						Improve learning		Parental education/literacy an obstacle for			
	Literacy		Numeracy		Communication		By setting time & space		Literacy		Numeracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Daughters											
Any Treatment (T1 or T2)	1.995*		1.683		1.150		1.551		-0.011		-0.016	
	(1.164)		(1.261)		(1.082)		(1.114)		(0.015)		(0.014)	
Videos (T1)		-0.629		-1.266		-1.810		-1.359		0.011		0.007
		(1.468)		(1.570)		(1.296)		(1.309)		(0.019)		(0.017)
Videos + Mobile apps (T2)		4.841***		4.882***		4.361***		4.707***		-0.035**		-0.040***
		(1.074)		(1.162)		(1.041)		(1.090)		(0.014)		(0.014)
Observations	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713	4,713
R-squared	0.197	0.207	0.203	0.215	0.187	0.201	0.192	0.203	0.259	0.261	0.264	0.266
Control Mean in F-up	91.883	91.883	91.030	91.030	92.506	92.506	91.507	91.507	0.131	0.131	0.129	0.129
T1 vs T2 (p-value)		0.000		0.000		0.000		0.000		0.013		0.007
	Sons											
Any Treatment (T1 or T2)	1.840*		1.582		0.339		0.418		-0.010		-0.009	
	(1.083)		(1.081)		(1.057)		(1.122)		(0.014)		(0.014)	
Videos (T1)		0.308		-0.419		-2.398*		-2.219		0.015		0.015
		(1.342)		(1.331)		(1.280)		(1.393)		(0.019)		(0.019)
Videos + Mobile apps (T2)		3.510***		3.762***		3.321***		3.292***		-0.037***		-0.036***
		(0.991)		(0.994)		(0.981)		(1.022)		(0.013)		(0.013)
Observations	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488	4,488
R-squared	0.193	0.196	0.213	0.219	0.189	0.201	0.202	0.212	0.260	0.263	0.271	0.274
Control Mean in F-up	92.130	92.130	91.652	91.652	92.553	92.553	91.842	91.842	0.127	0.127	0.122	0.122
T1 vs T2 (p-value)		0.004		0.000		0.000		0.000		0.007		0.009

Note: OLS regressions with clustered standard errors at the school catchment area level. The covariates considered in all regressions are: age of target child, gender of target child, education of household head (baseline), availability of food (baseline) and enumerator fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table A22: Statistical Balance at Baseline for *Sample 1 collected in March 2020*: Mean differences between treatment groups and control

Variables	T1 vs C	SE	P-value
	Socio-Demographic Characteristics		
Female	-0.025*	[0.01]	0.08
Age of target child	0.007	[0.06]	0.91
Age of father	0.837	[0.64]	0.20
Age of mother	0.049	[0.41]	0.90
Household size	0.024	[0.37]	0.95
Number of rooms	0.218	[0.15]	0.15
Always eat enough	-0.005	[0.03]	0.88
Not always eat enough	-0.004	[0.03]	0.91
Frequently not eat enough	0.009	[0.02]	0.68
HH education: none	-0.043	[0.05]	0.43
HH education: Islamiyya / other	0.011	[0.02]	0.62
HH education: any primary	0.003	[0.03]	0.91
HH education: any secondary	0.029	[0.02]	0.25
HH education: higher than secondary	0.000	[0.02]	0.99
Number of books	-0.019	[0.06]	0.76
More than 20 books	0.044	[0.05]	0.36
Less than 20 books	-0.025	[0.04]	0.59
Language spoken: Hausa	0.056	[0.06]	0.38
Language spoken: Fulani	-0.093*	[0.05]	0.07
Language spoken: Kanuri	0.007	[0.01]	0.33
Language spoken: other	0.030	[0.03]	0.37
Toilet: flush to piped sewer	-0.010	[0.01]	0.50
Toilet: pit latrine (VIP)	-0.005	[0.06]	0.93
Toilet: pit latrine w/slab	0.164*	[0.09]	0.09
Toilet: pit latrine without slab/open pit	0.022	[0.08]	0.79
Toilet: no facility/bush/field	-0.120	[0.10]	0.22
Toilet: other-ventilated, hanging and other	-0.051	[0.06]	0.42
Water: piped	-0.026	[0.03]	0.44
Water: public tap	0.011	[0.06]	0.84
Water: tubewell or borehole	0.028	[0.09]	0.77
Water: protected dug well	-0.004	[0.03]	0.89
Water: unprotected dug well	-0.014	[0.09]	0.88
Water: protected/unprotected spring	0.001	[0.00]	0.83
Water: surface water/river/canal/irrigation	0.005	[0.03]	0.85
Water: other source	-0.001	[0.00]	0.79
	Main Outcomes for Target Child: School Enrollment Uwezo, EGRA and EGMA		
Child currently attends formal school	-0.036	[0.07]	0.59
Letter recognition (EGRA)	0.203	[0.30]	0.50
Nonword reading (EGRA)	3.783	[3.29]	0.28
Oral passage reading (EGRA)	-0.101	[0.15]	0.51
Reading comprehension (EGRA)	0.047	[0.03]	0.14
Listening comprehension	-0.100	[0.20]	0.63
Letter dictation (EGRA)	0.261	[0.16]	0.12
Word dictation: correct word (EGRA)	0.031	[0.04]	0.44
Word dictation: all letters correct (EGRA)	0.033	[0.04]	0.37
Number recognition (EGMA)	0.020	[0.19]	0.91
Number discrimination (EGMA)	-0.429	[0.42]	0.31
Missing number (EGMA)	-0.039	[0.10]	0.71
Addition (EGMA)	-0.154	[0.16]	0.33
Substraction (EGMA)	-0.048	[0.10]	0.64

Note: HH refers to household head. OLS regressions with clustered standard errors at the school catchment area level. Dependent variable corresponds to each variable shown in this table at baseline on a dummy identifying T, T1 and T2, respectively. *** p<0.01, ** p<0.05, * p<0.1

Table A23: Statistical Balance at Baseline for *Sample 2 Pre- and Post-Covid collected in March 2020 & Sept-Dec with less than 10% of school closures*: Mean differences between treatment groups and control

Variables	T vs C	SE	P-value	T1 vs C	SE	P-value	T2 vs C	SE	P-value
	Socio-Demographic Characteristics								
Female	-0.010	[0.01]	0.25	0.019*	[0.01]	0.08	0.006	[0.01]	0.51
Age of target child	-0.049	[0.04]	0.19	0.039	[0.05]	0.45	0.054	[0.04]	0.20
Age of father	-0.149	[0.47]	0.75	-0.736	[0.65]	0.26	0.531	[0.48]	0.27
Age of mother	-0.198	[0.33]	0.55	-0.201	[0.44]	0.65	0.371	[0.36]	0.30
Household size	-0.113	[0.19]	0.56	0.215	[0.32]	0.50	0.069	[0.21]	0.74
Number of rooms	-0.089	[0.08]	0.29	-0.028	[0.13]	0.83	0.140	[0.09]	0.12
Always eat enough	-0.049**	[0.02]	0.04	0.046	[0.03]	0.12	0.050*	[0.03]	0.08
Not always eat enough	0.058*	[0.03]	0.07	-0.065	[0.04]	0.12	-0.056	[0.04]	0.12
Frequently not eat enough	-0.010	[0.02]	0.62	0.020	[0.03]	0.49	0.005	[0.02]	0.79
HH education: none	0.042	[0.04]	0.25	0.028	[0.05]	0.55	-0.073*	[0.04]	0.07
HH education: Islamiyya / other	-0.014	[0.01]	0.23	-0.002	[0.01]	0.89	0.021*	[0.01]	0.09
HH education: any primary	0.005	[0.02]	0.77	-0.001	[0.03]	0.95	-0.007	[0.02]	0.72
HH education: any secondary	-0.014	[0.02]	0.46	-0.012	[0.02]	0.56	0.025	[0.02]	0.25
HH education: higher than secondary	-0.020	[0.02]	0.28	-0.012	[0.02]	0.47	0.035	[0.02]	0.15
Number of books	0.032	[0.04]	0.39	-0.015	[0.06]	0.80	-0.039	[0.04]	0.32
More than 20 books	-0.005	[0.04]	0.89	-0.049	[0.04]	0.24	0.029	[0.04]	0.50
Less than 20 books	-0.026	[0.03]	0.36	0.064	[0.04]	0.12	0.010	[0.03]	0.75
Language spoken: Hausa	-0.052	[0.05]	0.25	0.034	[0.06]	0.58	0.060	[0.05]	0.21
Language spoken: Fulani	0.005	[0.03]	0.87	0.033	[0.05]	0.54	-0.021	[0.03]	0.45
Language spoken: Kanuri	0.036	[0.02]	0.11	-0.045**	[0.02]	0.04	-0.033	[0.02]	0.16
Language spoken: other	0.011	[0.02]	0.58	-0.022	[0.02]	0.18	-0.006	[0.02]	0.79
Toilet: flush to piped sewer	-0.007	[0.01]	0.22	0.001	[0.01]	0.88	0.010	[0.01]	0.18
Toilet: pit latrine (VIP)	-0.068*	[0.03]	0.05	0.060	[0.05]	0.21	0.072*	[0.04]	0.09
Toilet: pit latrine w/slab	-0.052	[0.06]	0.35	0.058	[0.09]	0.52	0.050	[0.06]	0.38
Toilet: pit latrine without slab/open pit	0.076	[0.07]	0.27	-0.181**	[0.09]	0.04	-0.030	[0.08]	0.69
Toilet: no facility/bush/field	0.089	[0.05]	0.10	-0.011	[0.10]	0.91	-0.123**	[0.05]	0.02
Toilet: other-ventilated, hanging and other	-0.037	[0.02]	0.13	0.073	[0.06]	0.23	0.022	[0.02]	0.31
Water: piped	0.012	[0.01]	0.35	-0.021*	[0.01]	0.09	-0.008	[0.02]	0.61
Water: public tap	0.043	[0.03]	0.22	-0.056	[0.05]	0.28	-0.038	[0.04]	0.32
Water: tubewell or borehole	0.053	[0.05]	0.32	-0.123	[0.08]	0.11	-0.022	[0.06]	0.70
Water: protected dug well	-0.030**	[0.01]	0.03	0.039	[0.02]	0.10	0.026*	[0.02]	0.09
Water: unprotected dug well	-0.084	[0.06]	0.14	0.147*	[0.08]	0.08	0.056	[0.06]	0.37
Water: protected/unprotected spring	0.002	[0.00]	0.41	0.002	[0.00]	0.63	-0.004*	[0.00]	0.08
Water: surface water/river/canal/irrigation	0.007	[0.01]	0.58	0.006	[0.02]	0.72	-0.013	[0.01]	0.31
Water: other source	-0.003	[0.00]	0.38	0.006	[0.01]	0.39	0.002	[0.00]	0.61
	Main Outcomes for Target Child: School Attendance, Uwezo, EGRA and EGMA								
Child currently attends formal school	0.006	[0.04]	0.88	-0.003	[0.05]	0.95	-0.007	[0.04]	0.86
Letter recognition (EGRA)	0.093	[0.17]	0.60	-0.179	[0.21]	0.39	-0.055	[0.20]	0.78
Nonword reading (EGRA)	2.228	[3.73]	0.56	-5.186	[3.33]	0.13	-0.058	[4.06]	0.99
Oral passage reading (EGRA)	-0.010	[0.08]	0.90	0.043	[0.07]	0.56	-0.004	[0.10]	0.97
Reading comprehension (EGRA)	0.001	[0.02]	0.98	-0.021	[0.01]	0.14	0.009	[0.03]	0.77
Listening comprehension	-0.038	[0.17]	0.82	-0.131	[0.19]	0.49	0.112	[0.19]	0.56
Letter dictation (EGRA)	0.073	[0.11]	0.53	-0.178	[0.14]	0.22	-0.027	[0.13]	0.84
Word dictation: correct word (EGRA)	0.003	[0.02]	0.89	-0.005	[0.03]	0.87	-0.002	[0.03]	0.93
Word dictation: all letters correct (EGRA)	-0.002	[0.02]	0.93	0.006	[0.03]	0.84	0.000	[0.03]	0.99
Number recognition (EGMA)	-0.042	[0.11]	0.70	-0.013	[0.14]	0.92	0.066	[0.13]	0.60
Number discrimination (EGMA)	-0.055	[0.26]	0.83	0.127	[0.38]	0.74	0.023	[0.27]	0.93
Missing number (EGMA)	-0.032	[0.06]	0.57	0.027	[0.09]	0.77	0.034	[0.06]	0.57
Addition (EGMA)	-0.081	[0.09]	0.34	0.073	[0.11]	0.52	0.084	[0.10]	0.39
Substraction (EGMA)	-0.041	[0.06]	0.50	0.027	[0.07]	0.70	0.047	[0.07]	0.52

Note: HH refers to household head. OLS regressions with clustered standard errors at the school catchment area level. Dependent variable corresponds to each variable shown in this table at baseline on a dummy identifying T, T1 and T2, respectively. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A24: Statistical Balance at Baseline for *Sample 3 Pre- and Post-Covid collected in March and Nov-Dec 2020 (most of the schools opened)*: Mean differences between treatment groups and control

Variables	T vs C	SE	P-value	T1 vs C	SE	P-value	T2 vs C	SE	P-value
				Socio-Demographic Characteristics					
Female	-0.004	[0.01]	0.68	0.012	[0.01]	0.26	-0.001	[0.01]	0.90
Age of target child	-0.054	[0.04]	0.14	0.020	[0.05]	0.70	0.078*	[0.04]	0.06
Age of father	-0.031	[0.45]	0.94	-0.808	[0.58]	0.16	0.604	[0.48]	0.21
Age of mother	-0.078	[0.33]	0.81	-0.378	[0.41]	0.35	0.392	[0.38]	0.30
Household size	-0.111	[0.20]	0.58	0.206	[0.28]	0.46	0.045	[0.22]	0.84
Number of rooms	-0.047	[0.08]	0.57	-0.026	[0.12]	0.82	0.098	[0.09]	0.30
Always eat enough	-0.029	[0.02]	0.22	0.035	[0.03]	0.20	0.025	[0.03]	0.41
Not always eat enough	0.022	[0.03]	0.46	-0.037	[0.04]	0.33	-0.012	[0.04]	0.73
Frequently not eat enough	0.007	[0.02]	0.73	0.002	[0.03]	0.92	-0.013	[0.02]	0.52
HH education: none	0.031	[0.03]	0.36	0.007	[0.04]	0.86	-0.058	[0.04]	0.15
HH education: Islamiyya / other	-0.018	[0.01]	0.11	0.015	[0.02]	0.34	0.021*	[0.01]	0.08
HH education: any primary	0.012	[0.02]	0.54	-0.020	[0.02]	0.41	-0.006	[0.02]	0.76
HH education: any secondary	-0.004	[0.02]	0.78	0.001	[0.02]	0.95	0.006	[0.02]	0.73
HH education: higher than secondary	-0.020	[0.02]	0.27	-0.003	[0.01]	0.82	0.036	[0.03]	0.17
Number of books	0.027	[0.04]	0.44	-0.023	[0.05]	0.65	-0.03	[0.04]	0.43
More than 20 books	0.001	[0.03]	0.99	-0.023	[0.04]	0.55	0.015	[0.04]	0.72
Less than 20 books	-0.028	[0.03]	0.29	0.046	[0.03]	0.19	0.016	[0.03]	0.60
Language spoken: Hausa	-0.043	[0.05]	0.35	0.028	[0.05]	0.61	0.053	[0.05]	0.30
Language spoken: Fulani	-0.001	[0.03]	0.96	0.037	[0.05]	0.42	-0.023	[0.03]	0.40
Language spoken: Kanuri	0.035	[0.02]	0.12	-0.043**	[0.02]	0.05	-0.029	[0.02]	0.22
Language spoken: other	0.009	[0.02]	0.64	-0.022	[0.02]	0.17	-0.001	[0.03]	0.98
Toilet: flush to piped sewer	-0.009	[0.01]	0.22	0.010	[0.01]	0.37	0.008	[0.01]	0.34
Toilet: pit latrine (VIP)	-0.058*	[0.03]	0.08	0.071*	[0.04]	0.10	0.050	[0.04]	0.22
Toilet: pit latrine w/slab	-0.052	[0.06]	0.35	0.058	[0.08]	0.46	0.048	[0.06]	0.41
Toilet: pit latrine without slab/open pit	0.079	[0.07]	0.24	-0.181**	[0.08]	0.02	-0.008	[0.08]	0.92
Toilet: no facility/bush/field	0.084	[0.05]	0.12	-0.022	[0.08]	0.80	0.126**	[0.05]	0.01
Toilet: other-ventilated, hanging and other	-0.043	[0.03]	0.10	0.064	[0.05]	0.20	0.029	[0.03]	0.26
Water: piped	-0.004	[0.02]	0.78	0.012	[0.03]	0.66	-0.001	[0.02]	0.97
Water: public tap	0.027	[0.03]	0.41	-0.035	[0.04]	0.44	-0.022	[0.04]	0.56
Water: tubewell or borehole	0.027	[0.05]	0.61	-0.083	[0.07]	0.24	0.012	[0.06]	0.83
Water: protected dug well	-0.017	[0.01]	0.21	0.026	[0.02]	0.20	0.012	[0.02]	0.45
Water: unprotected dug well	-0.051	[0.05]	0.35	0.085	[0.07]	0.25	0.027	[0.06]	0.64
Water: protected/unprotected spring	0.002	[0.00]	0.49	0.001	[0.00]	0.72	-0.004*	[0.00]	0.09
Water: surface water/river/canal/irrigation	0.018	[0.02]	0.26	-0.009	[0.02]	0.62	-0.024	[0.02]	0.12
Water: other source	-0.001	[0.00]	0.67	0.004	[0.01]	0.52	-0.000	[0.00]	0.95
				Main Outcomes for Target Child: School Attendance, Uwezo, EGRA and EGMA					
Child currently attends formal school	0.006	[0.04]	0.88	0.005	[0.05]	0.91	-0.013	[0.04]	0.76
Letter recognition (EGRA)	0.094	[0.17]	0.59	-0.090	[0.20]	0.66	-0.097	[0.21]	0.64
Nonword reading (EGRA)	1.662	[3.47]	0.64	-3.847	[3.05]	0.22	0.188	[4.07]	0.96
Oral passage reading (EGRA)	-0.094	[0.07]	0.17	0.092	[0.09]	0.28	0.096	[0.07]	0.20
Reading comprehension (EGRA)	0.020	[0.01]	0.15	-0.018	[0.01]	0.22	-0.021	[0.01]	0.13
Listening comprehension	-0.072	[0.16]	0.66	-0.022	[0.17]	0.90	0.137	[0.19]	0.48
Letter dictation (EGRA)	0.099	[0.10]	0.34	-0.108	[0.13]	0.40	-0.094	[0.12]	0.45
Word dictation: correct word (EGRA)	-0.005	[0.02]	0.81	0.010	[0.03]	0.74	0.002	[0.03]	0.94
Word dictation: all letters correct (EGRA)	-0.006	[0.02]	0.77	0.013	[0.03]	0.62	0.001	[0.03]	0.96
Number recognition (EGMA)	-0.053	[0.11]	0.62	0.063	[0.13]	0.64	0.045	[0.13]	0.73
Number discrimination (EGMA)	-0.148	[0.25]	0.55	0.175	[0.34]	0.61	0.130	[0.26]	0.61
Missing number (EGMA)	-0.045	[0.06]	0.42	0.051	[0.08]	0.53	0.041	[0.06]	0.51
Addition (EGMA)	-0.147*	[0.09]	0.09	0.170	[0.12]	0.14	0.131	[0.11]	0.22
Substraction (EGMA)	-0.087	[0.06]	0.16	0.081	[0.07]	0.25	0.092	[0.08]	0.26

Note: HH refers to household head. OLS regressions with clustered standard errors at the school catchment area level. Dependent variable corresponds to each variable shown in this table at baseline on a dummy identifying T, T1 and T2, respectively. *** $p<0.01$, ** $p<0.05$, * $p<0.1$