Role of EdTech on numeracy and literacy for children with functional difficulties: Experimental evidence from Kenya

Frank Odhiambo * Isabel Günther †

February 15, 2025

Abstract

We conduct a field experiment to investigate the effect of an educational technology intervention on the numeracy and literacy skills of children with functional difficulties in a low-income setting with a high disability prevalence. Children with special needs in primary school are recruited through a government screening program at school. After randomization, at the school level, children in the intervention group are offered a computer-assisted-learning mobile app. Eight months post-intervention, we find that the intervention has positive and significant effects on literacy rates, with effect sizes exceeding the median effects reported in prior studies exploring literacy interventions. However, we find no significant effects on numeracy rates. Our disaggregated results shows that the intervention improves lower-order literacy skills and higher-order numeracy skills. On literacy skills, where aggregate effects are positive and significant, both males and females benefit from the intervention. Treatment effects are primarily driven by treatment compliance—whether students accept and use the tablet—rather than by changes in study effort, time allocation, perceptions about schooling, interactions with teachers and peers or in mental well-being. Our findings suggest that EdTech intervention may be effective in improving learning outcomes and that effectiveness is dependent on compliance with the treatment rather than ancillary behavioral or psychological mechanisms.

Keywords: EdTech, literacy, numeracy, disability

*Development Economics Group, ETH Zurich, Switzerland. E-mail: fodhiambo@ethz.ch.

†Development Economics Group, ETH Zurich, Switzerland. E-mail: isabel.guenther@nadel.ethz.ch.

1 Introduction

Universally, school enrollment has increased over the years (Glewwe and Kremer, 2006; World Bank, 2018) such that by 2010, gross enrollment rate was above 100 percent in Sub-Saharan Africa and South East Asia, regions that previously had much lower school enrollment rates (World Bank, 2018). More recently, educational efforts are more increasingly shifting toward learning outcomes (Glewwe et al., 2017). Despite widespread school attendance, evidence indicates a significant learning gap in various regions with many children demonstrating numeracy and literacy skills that lag several grades behind their current educational level (World Bank, 2018).

One particular demographic that is particularly disadvantaged is children with functional difficulties. For example, Zhang and Holden (2023), find that those with physical and intellectual difficulties have significantly lower numeracy skills. For children with disabilities, there are several reasons why their numeracy and literacy levels might be even more severely impacted. First, due to their special needs or disabilities, they are much less likely to attend school regularly (Zhang and Holden, 2023). Access to school may be a barrier if a child with mobility difficulties has to travel several kilometres to school (Trani et al., 2012), or if they do not have access to assistive devices such as wheelchairs (Lamichhane, 2013), both problems which are particularly pertinent in developing countries.

Moreover, in numerous instances, children with disabilities are enrolled in mainstream schools rather than specialized institutions, largely due to the limited availability of the latter (McKinney and Swartz, 2016). These mainstream schools are typically less equipped with teachers with specialized training in addressing the unique needs of students with disabilities. For instance, children with hearing impairments often require instructional adaptations, such as slower speech and unobstructed visibility of the speaker's face, to facilitate effective lip reading (Lockwood, 2006). With conventional education environments failing to accommodate the diverse learning requirements of students with disabilities (Yuwono and Okech, 2021), such systemic shortcomings may have significant implications for the academic and developmental outcomes of children with functional difficulties, particularly when they require tailored support to successfully integrate.

How can learning outcomes for children with disabilities be improved? The majority of the existing literature on interventions and policies aimed at enhancing learning outcomes for disabled students in low-income settings emphasizes environmental factors, such as improving physical access and providing technical support (Hanafin et al., 2007), as well as fostering social support mechanisms to promote inclusion (Moriña Díez, 2010).

Specific studies have investigated the impact of assistive devices. For example, a study in rural China finds that wearing eyeglasses can significantly increase test scores and lead to additional years of schooling, particularly benefiting underperforming students (Glewwe et al. (2016)). However, they also note gender disparities, with girls being less likely to wear glasses, and thus less likely to benefit from such interventions. Further, Ma et al. (2014) find that providing free eyeglasses has a greater impact on academic performance than parental education or family wealth. Wang et al. (2017) examine the effects of providing free eyeglasses on subsequent purchases of glasses by individuals with visual impairments while Grimm and Hartwig (2018) analyze willingness to pay for eyeglasses in Burkina Faso. Research

Beyond vision-related interventions, Harris and Terlektsi (2011) explore the comparative effects of hearing aids and cochlear implants on the reading and spelling abilities of deaf adolescents. While both groups lag behind chronological age-level benchmarks, students using hearing aids perform better than those with cochlear implants, highlighting the potential of assistive devices to mitigate educational challenges faced by children with disabilities.

To date, there are no causal inference studies that evaluate the impact of educational technology (EdTech) on the educational attainment of children with disabilities, particularly in low-income settings. Education technology holds particular promise for several reasons. First, children with functional difficulties can benefit from digitally delivered learning materials. These resources allow students to progress at their own pace, making education accessible even for those with irregular school attendance. This flexibility is vital in addressing constraints on school access, as technology can enable learning regardless of physical barriers. In high-income contexts, digital devices are increasingly used to support specialized learning needs but scale-up is low given that a higher teacher-to-student ratio in specialized schools already supports special learning needs.(Bastawrous and Armstrong, 2013). In low-income settings, take-up is still low.

Second, recent studies highlights the effectiveness of "teaching at the right level" (Banergee et al., 2016), an approach that tailors instruction to students' actual abilities rather than adhering to a uniform curriculum. This strategy is particularly relevant in low-income settings, where classrooms often include students with vastly different learning levels. For children with disabilities, the challenge of differentiated instruction is even greater, as their needs often fall outside the scope of traditional teaching methods. EdTech, by offering personalized and adaptive learning pathways, is uniquely suited to address these gaps, potentially improving learning outcomes for children with disabilities in low-income settings.

Outside of disability status, the evidence on the impact of EdTech regarding literacy and numeracy outcomes is mixed, but largely promising. While some studies highlight significant benefits, others indicate limited or no impact. For instance, Rodriguez-Segura (2022) finds that technology-based interventions focused on self-regulated study have the greatest potential for improving learning outcomes. In Botswana, Angrist et al. (2022)) find that SMS and phone calls are cost-effective solutions for improving learning outcomes while in El Salvador, Büchel et al. (2022) find that classes with computer-assisted learning have higher numeracy gains than those with additional teacher-led instruction. Agrawal et al. (2022) suggest adding personalized content recommendations to increase EdTech usage, with the greatest learning benefits accruing for more frequent users with more personalized content interactions. Conversely, Piper et al. (2016) finds that EdTech fails to yield measurable gains. They also suggest the need for context-specific interventions, particularly in low-resource settings where factors like access to devices, teacher capacity, and digital literacy may significantly influence the effectiveness of EdTech interventions. Overall, Escueta et al. (2020) conclude that while providing access to technology by itself may not generate large learning gains, computer-assisted-learning (CAL) and technology-enabled behavioral interventions demonstrate the largest potential.

In this paper, we provide experimental evidence on the role of educational technology in improving learning outcomes for a population that has not yet been empirically studied–children with functional difficulties. We study the effect of CAL on the numeracy and literacy of children with functional difficulties in a largely rural setting in Kenya, with a high disability prevalence. Specifically, we recruit 624 children with functional difficulties in Grades 3, 4 and 5 across 63 primary schools in Homabay County. After randomly assigning treatment status at school level, we offer the treated sample a tablet device with a learning software installed on it. Eight months post-intervention, we measure treatment effects on our numeracy and literacy outcomes, based on a standardized test administered to the students.

We find that the EdTech intervention positively impacts literacy among children with functional difficulties, with effect sizes exceeding the median reported in studies examining various interventions related to literacy, but find no significant effects for numeracy. Our disaggregated literacy and numeracy analysis shows significant treatment effects in lower-order literacy skills and higher-order numeracy skills. These treatment effects are predominantly explained by compliance with the intervention—acceptance and use of the tablet—rather than changes in study effort, study time, perceptions of schooling, or interactions with teachers and peers. Furthermore, we find no evidence that changes in mental health or well-being explain the observed treatment effects. For literacy, where our main results reveal aggregate positive effects, both male and female students benefit from the intervention. However, the effect size for females is smaller–but statistically insignificant. Nevertheless, females have significantly smaller gains in higher-order literacy and numeracy skills compared to their male counterparts.

The rest of the paper is set up as follows. In Section 2, we discuss study design, setting and sample selection, in Section 3 we discuss the intervention while in Section 4 we discuss the data and empirical strategy. In Section 5, we show the background characteristics of our participants and then in Section 6, we discuss the results of the experiment. Finally in Section 7, we make some conclusions and recommendations.

2 Design, setting and sample

2.1 Design and study setting

To evaluate the role of EdTech on learning outcomes, we implemented a cluster-randomized experiment with randomization at the school level. Sampled schools[1] were randomized to a treatment or control group, with the treatment sample receiving a tablet pre-installed with an educational device, together with a solar panel for charging the tablet, and the control group receiving neither[2]. Randomization was done at school level to minimize the likelihood of spillovers.

We implemented our study in Homabay County, a largely rural setting, in Western Kenya. In 2019, the over-5 disability prevalence in Homabay County was twice the national average (4.3% vs 2.2%- Kenya National Bureau of Statistics (2020)), making it an ideal location to carry out our experiment. In 2019, the county had 1,042 primary schools. Of these, 85% were public schools (881), hosting just about 90% of all children enrolled in primary school in the county (Government of Kenya, 2019). For this reason, in our experiment, we focused only on children attending the 881 public primary schools.

The study was registered with the American Economic Association's (AEA) registry (Odhiambo and Gunther, 2023). Ethical approval was obtained from the ETH Zurich Ethics Committee[3] and the Strathmore University Institutional Scientific and Ethical Review Committee[4] in Kenya. Furthermore, the research activities in Kenya were authorized by the National Commission for Science, Technology, and Innovation (NACOSTI)[5], ensuring adherence to

[1]See Section 2.2 for more details on the sample selection process.
[2]See Section 3 for more details on the intervention.
[3]Approval number: EK 2023-N-11
[4]Approved as SU-ISERC1564/23 and renewed as SU-ISERC1975/24
[5]License No: NACOSTI/P/23/24649

local regulatory and ethical guidelines.

2.2 Sample and sample recruitment

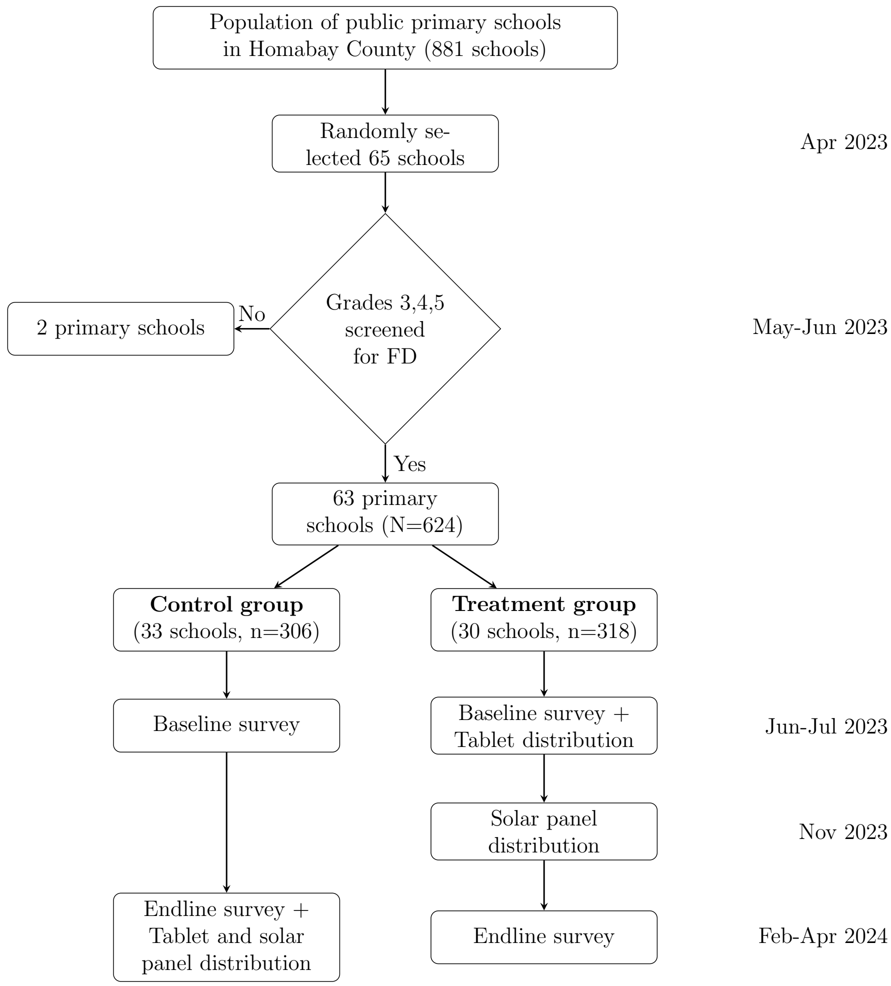
We recruited school-going or recently dropped-out children who were positively screened for functional difficulties and who were in their 3rd, 4th or 5th grade. First, we randomly selected sixty-five (65) out of 881 public schools in the county[6]. Two out of the sixty-five schools could not be reached and were consequently excluded from the study. One of the schools was located in an area that at the time had a cholera outbreak, thus posing health risks to our team of enumerators. The other was located in an area that had significant travel and logistical difficulties for the assessment officers.

Next, we used the support of a government agency responsible for disability screening in schools, the County Disability Assessment Office (CDAO), to identify our sample within these schools. Within each sampled school, CDAO officers scheduled a disability screening date and notified its head. Before this screening date, the CDAO officers asked the teachers of Grades 3, 4, and 5 to pre-identify children in their classes whom they suspected of having any form of disability or functional difficulty. The teachers would then invite these children and their parents or caregivers for a disability screening exercise at their respective schools on the communicated date for further evaluation by the screening officers. During the disability screening, the CDAO officers received the list of the pre-identified children from the teachers and then conducted a one-on-one meeting with each one, in the presence of their parent or caregiver. For each child, the CDAO discussed with the parent or caregiver, to understand the underlying reason for the child being pre-identified for screening. They then used a set of government approved screening forms and equipment to verify whether the child had a functional difficulty. Following this assessment, the CDAO officers either confirmed the child as having a functional difficulty referred them for more specialist assessment or ruled out the existence of one. Children were typically referred for more specialist assessment if the CDAO did not have the requisite equipment on-site or if they recommended medical intervention.

At the screening, the CDAOs notified the attending parents or caregivers about our study and sought consent to have them share the screening data with us for research purposes. The CDAO shared with the research team the screening data for all consenting caregivers. Demand for FD screening was high in the sampled schools. Even though the CDAO officers communicated to the school that the screening would target only children in Grades 3, 4

[6] In a pilot exercise conducted in 10 schools in February 2023, we determined that there were, on average, 2.7 children with functional difficulties per grade in Grades 3 to 5. This average was used to estimate the number of schools needed to achieve the desired sample size for the study.

Figure 1: Research Design and timeline



Notes: Figure shows the timeline of the experiment. Number of public schools in the county is based on the Kenya Basic Education Statistical Booklet (Government of Kenya, 2019). Two of the 65 sampled schools could not be reached for FD screening due to logistical and disease outbreak reasons and were therefore dropped from the study. Four months after the intervention, we distributed solar panels to address a tablet charging due to electricity access challenges. Details of the intervention are discussed in Section 3.

and 5, children from across all Grades in the sampled primary school attended (See Table A2 in Appendix). Nevertheless, most of the children who attended the screening were from our target Grades. Most of the children who were screened by the CDAOs were confirmed to have a FD (88%) while another 8% were referred for more specialized assessment. Only 4% of those who attended the screening were not confirmed to have a FD. Separately, about 70% of those screened were within our target Grades. Only children with FD or referred for more specialist screening in Grades 3,4 and 5 were selected for the study.

After the FD screening was completed, we then randomly assigned each of the 63 screened schools to one of two groups: a treatment group, in which study participants received the intervention, and a control group, in which participants did not. A total of 30 schools were assigned to the treatment group with the remaining 33 being assigned control group (See Figure 1).

3 Intervention

The intervention that was administered in this study is a self-study mobile application (mobile app) installed on a low-cost tablet device. Children with functional difficulties in the treatment group were offered an 8-inch android-based tablet (Figure A1) and a solar panel (Figure A2) both sourced locally in the country. The tablets were pre-installed with ANTON, a self-regulated educational technology (EdTech) application offering curriculum content in various subjects, including Mathematics, English, Science, and Music. ANTON also provides resources in Geography, Physics, Biology, and several additional languages such as German, French, Italian, Portuguese, and Spanish. The app is freely available for download and use. We show examples of some of the content available in the app in the Appendix A3.

ANTON incorporates a unique token-based gaming system designed to incentivize learning. For each educational activity a user completes successfully, they earn coins which can then be used to play various games within the app. This gamification aspect is designed to enhance user engagement and motivation to stay engaged. We procured from ANTON, a subscription package, that offers an offline access feature, enabling users to download and thereafter access learning materials without needing an internet connection.

Before distributing the tablets to the participants, they were pre-set to ensure only the necessary applications and functionalities were active. This involved uninstalling any non-essential software, deactivating the Play Store, and removing potentially distracting apps such as YouTube, Maps, and WhatsApp. We retained essential operational features such as network connectivity, SIM slot functionality, messaging, and calling capabilities.

The tablets were issued to the children immediately after administering the baseline survey questionnaire. Tablets were distributed at the child's residential place and in the presence of a caregiver, who had to consent to the child receiving the device. Before issuing the tablets, the research team conducted brief demonstration sessions, lasting about ten minutes each, for the children and their parents or caregivers. These sessions provided instructions on using the tablet and the ANTON app, including how to switch subjects or levels and access the games. Both the caregiver and the child were informed that the tablet was intended to supplement, not replace, teacher efforts in school and that it was recommended for home use rather than school use. Although the research team communicated this recommendation, there were no mechanisms to enforce it, leaving the possibility that some children might use the tablets in school.

Three months after the tablets were issued, even though most households reported having access to solar electricity, several participants indicated that the tablets were rapidly depleting their existing solar panel resources. This depletion significantly limited their access to lighting at night and therefore several households were no longer charging the tablets at home or at all. In response to this issue, the research team procured low-cost solar panels specifically designed to handle the power requirements of the tablets (costing about USD 18 each) to ensure that households could continued using the tablets without compromising their essential lighting or other energy needs.

4 Data and Empirical Strategy

4.1 Data

We combined assessment data from the CDAO and field survey data for our experiment, collecting baseline and endline surveys. For both the baseline and endline surveys, we administered three types of interviews: a caregiver survey, a child survey, and an intervention survey. In the caregiver survey, we collected demographic and household information, including household size, wealth indicators, technology adoption, and the education levels of parents and siblings. We also collected information on the study child's education, well-being, school participation, and any functional difficulties. At endline, caregivers were asked additional questions about their usage of the tablets. In the child survey, we captured details on school attendance, mental health, and well-being. At endline, we also collected a similar tablet usage module for the treated sample survey, similar to the one administered to theparent, to compare usage reports between the child and the parent. Lastly, our intervention survey, administered at baseline, gathered information related to the issuance of the tablets, including consent for the child to participate, details of the device, and the timing of its distribution to the intervention group.

4.1.1 Literacy and numeracy tests for siblings

Besides our sampled children, we also administered the numeracy and literacy tests to a sibling, applying a specific inclusion criteria for the sibling to test. We adminsitered the test to a sibling of the sampled child only if they were between 5–17 years old and had no more than a primary school level of education. If multiple siblings met these criteria, we prioritized interviewing the younger sibling, ideally one younger than the study child. However, if the younger sibling was unavailable at the time of the interview, we selected the older sibling who still met our overall criteria.

4.2 Empirical Strategy

We use the following intention to treat (ITT) estimation equation:

$$y_i = \gamma_0 + \gamma_1 Treat_i + \gamma_2 BaselineScore_i + \epsilon_i \tag{1}$$

where y_i represents our outcome of interest, represented by the endline test score for Student i, γ_1 is the treatment effect estimate, and γ_2 the baseline score for Student i. We cluster our standard errors at the school level, which is also our level of randomization.

5 Background characteristics and balance table

Table 1, our balance table, provides details of the treatment and control sample study participants' baseline characteristics ex-ante. In the first part, we aggregate the various FDs the children were identified with into five major disability domains. The most dominant FD at the sampled schools was Learning FD for which 59% of our sample was positively screened. About 20% had Visual FD, another 19% Auditory FD while about 9% was positively assessed for Physical FD, those impairing mobility and other physical activity. About 8% had behaviour-related FDs[7].

Just about 42% of our sample were female and the average age of the participants was approximately 11.36 years, coming from households whose poverty likelihood was 67%. On average, they had attended school just under 4 days the previous week, and spent just about 45 minutes studying at home the previous week. Regarding our primary outcomes of interest, approximately 60% of participants could read a simple phrase. However, on average, a child could read only about 30 words from a 72-word paragraph. Comprehension skills were also quite low, with children correctly answering only 28% of the comprehension questions related to the passage on average. Numeracy skills were also particularly low, to begin with, and

[7]This functional difficulty categorization reflects the domains as aggregated by the research team, based on the categories used by the County Disability Assessment Office, which had some overlapping areas.

significantly diminished with increased difficulty. For example, while on average our sample was able to identify 63% of number shown to them, they could only get correctly on average just over 30% of division-related questions. Overall, our balance table shows balance between the treatment and control samples at baseline.

Table 1: Balance Table

Variable	**Control**	**Treated**	**p-value**
Functional Difficulty Domain			
Vision	0.18	0.22	0.20
Auditory	0.22	0.16	0.08
Physical	0.08	0.09	0.65
Learning	0.57	0.60	0.61
Social and Behaviour	0.07	0.09	0.44
Background			
Female	0.43	0.41	0.52
Age	11.35	11.34	0.94
Poverty likelihood	0.67	0.67	0.58
Schooling			
Grade (school level)	3.91	3.97	0.48
Days attended school last week	3.84	4.11	0.07
Hours studying at home yesterday	0.75	0.77	0.78
Literacy			
Can read simple phrase	0.60	0.60	0.93
Number of words read correctly (out of 72)	30.05	29.58	0.86
No. of comprehension questions correct (out of 5)	1.29	1.49	0.21
Numeracy			
Number identification (out of 4)	2.47	2.57	0.32
Which number is bigger (out of 3)	2.31	2.41	0.21
Addition (out of 4)	1.99	1.99	0.98
Skip pattern (out of 3)	1.39	1.43	0.60
Multiplication (out of 4)	1.24	1.44	0.11
Division (out of 3)	0.88	1.02	0.17
Mental well-being			
Positive emotional state (range 5-30)	22.73	22.80	0.84
Positive outlook (range 5-30)	21.65	21.88	0.56
Social desirability (range 3-15)	11.14	11.37	0.29
N	270.00	295.00	NA

Notes: Table shows the balance table of our treated and control samples at baseline. The assessment officers reported 13 FD categories, which we consolidate into 5 FD domains. The specific literacy and numeracy test questions administered to the participants are shown in Appendix A.8 and Appendix A.9 respectively. The mental well-being measures are based on likert scale questions each with 1-5 responses (lowest to highest).

5.1 Attrition

Our overall attrition rate at endline is 3% for the child survey and 5% for the parent survey. While attrition is slightly higher in the control sample (5% and 6% respectively) compared to the treatment sample (3% and 5% respectively), this difference is not statistically significant (Table A1).

6 Results

6.1 Effects on literacy and numeracy outcomes

6.1.1 Average treatment effects

We present the average treatment effect results in Table 2. Models 1–3 display literacy outcomes, while Models 4–6 present numeracy outcomes under varying specifications. In Models 1 and 4, treatment effects are estimated without any controls, though standard errors are clustered at the school level. Models 2 and 5 incorporate controls for students' baseline scores, while Models 3 and 6 add further covariates, including child age, gender, school grade, female household head's years of education, and household poverty likelihood. The results across Models 2, 3, 5, and 6 indicate that our findings are robust to the exclusion of child- and household-specific controls. For the remainder of the paper, we base our analysis and discussion on the specifications in Models 2 and 5, as outlined in our empirical strategy. (Section 4.2).

We find positive effects of the technological intervention on literacy with the treatment increasing literacy scores by 0.15 standard deviations, at the 90% confidence level. To contextualize our findings, the effect sizes we observe for literacy outcomes surpass the median effect sizes reported in most field experiments that measure literacy improvements (Q1 = 0.03, Median = 0.14, Q3 = 0.32; Evans and Yuan, 2022). Even though we do not find significant effects on numeracy, our standard errors indicate the effects would be significant at the 85% confidence level, which may be driven by sample size, power and the variability in our cluster sizes[8].

In Table 3, we present our results of the effects of the treatment across disaggregated literacy and numeracy measures, offering nuanced insights into how the intervention influences various aspects of learning outcomes. Our first literacy outcome, "Read," indicates whether the child could read a simple phrase, while outcomes Q1-Q2 measure whether they could answer comprehension questions related to the simple phrase. We categorize these as lower-order literacy skills. The third literacy outcome, "Words," reflects the number of words a child could read correctly from a 72-word passage. Outcomes Q3-Q7 assess how many of five com-

[8]We show in Figure A4 that there is large variability in the number of children between schools in our sample.

Table 2: Effect of EdTech on numeracy and literacy

	Literacy			Numeracy		
	1	2	3	4	5	6
Treat	0.182	0.153*	0.149*	0.148	0.098	0.090
	(0.126)	(0.079)	(0.081)	(0.098)	(0.065)	(0.066)
Female			0.187**			-0.048
			(0.078)			(0.058)
Age			-0.039*			-0.040**
			(0.022)			(0.016)
HH poverty likelihood			0.108			0.515
			(0.337)			(0.337)
Female HHH Ed. years			0.001			0.021*
			(0.012)			(0.012)
Grade			0.092*			0.116**
			(0.048)			(0.046)
Baseline score		0.617***	0.591***		0.610***	0.573***
		(0.038)	(0.043)		(0.036)	(0.040)
Clustered SEs (school)	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	565	565	560	565	565	560
N Clusters	63	63	63	63	63	63

Notes: Table shows the average treatment effect (ATE) of our EdTech intervention on learning outcomes. Both numeracy and literacy outcomes are standardized z-scores, which combines several observed measures from a standardized test. The literacy index includes: whether the child can read a simple phrase; number of correct questions related to the simple phrase (out of 2); the number of words the child can read from a 72-word passage; and number of correctly answered questions related to the longer passage (out of 5). The numeracy index comprises the number of correct responses to: four number identification questions; three questions on which number is bigger; four addition questions; three skip-pattern questions; four multiplication questions; and four division questions. Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

prehension questions related to this passage the child could answer correctly, representing higher-order literacy skills (See Appendix A.8, for specific questions).

On numeracy, the first outcome, "Num," measures how many numbers—out of four—a child could correctly identify. The "Big" outcome measures the ability to correctly identify the larger number in three pairs. These are categorized as lower-order cognitive numeracy skills. Conversely, "Add" measures the number of four addition problems solved correctly, and "Skip" evaluates the number of four number sequence problems correctly completed. The outcomes "Mult" and "Div" assess performance on four multiplication and division questions each. These latter four measures are considered higher-order cognitive numeracy skills (See Appendix A.9, for numeracy specific questions). Our notion of higher-versus lower-order cognitive skills is backed by the fact that we observe that the proportion of children correctly answering questions decreases with increasing difficulty, which is evident in the sample mean

Table 3: Treatment effect by dissagregated numeracy and literacy measures

	Literacy					Numeracy					
	Read	Q1	Q2	Words	Q3-Q7	Num	Big	Add	Skip	Mult	Div
Treat	0.078^{**}	0.095^{**}	0.076^{**}	4.602	0.259	0.072	-0.014	0.225^{*}	0.118	0.334^{**}	0.258^{*}
	(0.039)	(0.039)	(0.038)	(2.814)	(0.202)	(0.113)	(0.088)	(0.122)	(0.126)	(0.131)	(0.152)
Baseline score	0.507^{***}	0.477^{***}	0.475^{***}	0.576^{***}	0.634^{***}	0.472^{***}	0.410^{***}	0.477^{***}	0.680^{***}	0.486^{***}	0.547^{***}
	(0.045)	(0.039)	(0.046)	(0.040)	(0.039)	(0.048)	(0.050)	(0.038)	(0.047)	(0.043)	(0.048)
Ctrl End Mean	0.616	0.559	0.573	36.197	1.785	2.975	2.355	2.179	1.867	1.523	1.237
Controls	No	No	No	No	No	No	No	No	No	No	No
Clustered SEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	565	565	565	565	565	565	565	565	565	565	565
N Clusters	63	63	63	63	63	63	63	63	63	63	63

Notes: Robust standard errors in parenthesis with the following significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. The literacy measures include: a binary indicator for whether the child can read a simple phrase; whether they correctly answer two comprehension questions related to the simple phrase; the number of words the child can read from a 72-word passage; and a binary indicator for whether they correctly answer five comprehension questions related to the longer passage. The numeracy measures comprises: the number of correct responses to four number identification questions; the number of correct responses to three questions on which number is bigger; the number of addition questions (out of four) answered correctly; the number of correct responses to three skip-pattern questions; the number of correct answers to four multiplication questions; and the number of correct responses to three division questions. We cluster our standard errors at the school level.

proportions for the different question categories (See the control sample endline mean scores in Table 3).

Our results suggest that our participants' numeracy and literacy skills may respond differently to EdTech interventions. We find that the technological intervention has an effect only on; lower-order literacy skills and; higher-order cognitive numeracy skills. On literacy, we see effects on their ability to read the simple phrase and answer its associated questions–treated children are 8 percentage points (pp) more likely to read the simple phrase, and 9pp and 7pp more likely to correctly answer the two questions related to the passage– but no effect on the longer passage or its associated questions. On numeracy, we find no significant effect on lower-order cognitive skills, but a significant effect on two of four higher order skills—a 10% and 21% change in addition and multiplication competencies respectively following the EdTech intervention.

6.1.2 Heterogeneity by gender

Next, we examine gender heterogeneity in the impact of EdTech on our sample to determine whether the technology affects learning outcomes differently for boys and girls with FD, which we show in Table 4. Regarding literacy, where we reported significant aggregate positive effects in our main results, we find that males and females both benefit from our EdTech intervention. Even though the magnitude of the effect is smaller for females, this difference is statistically insignificant. We also observe that females with FD in the control sample have, on average, significantly higher literacy skills than their male counterparts, suggesting pre-existing gender differences in literacy rates. There is no significant gender differences in numeracy skills between these groups.

Table 4: Treatment heterogeneity by gender

	Literacy	Numeracy
Treat x Female	-0.123	-0.154
	(0.143)	(0.114)
Treat	0.209^{**}	0.161^{*}
	(0.101)	(0.088)
Female	0.260^{**}	0.038
	(0.109)	(0.089)
Baseline score	0.607^{***}	0.610^{***}
	(0.039)	(0.035)
Clustered SEs	Yes	Yes
Num. obs.	565	565
N Clusters	63	63

Notes: Table shows treatment effects by gender. Both numeracy and literacy outcomes are standardized z-scores, which combines several observed measures from a standardized test. The literacy index includes: whether the child can read a simple phrase; number of correct questions related to the simple phrase (out of 2); the number of words the child can read from a 72-word passage; and number of correctly answered questions related to the longer passage (out of 5). The numeracy index comprises the number of correct responses to: four number identification questions; three questions on which number is bigger; four addition questions; three skip-pattern questions; four multiplication questions; and four division questions. Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

When disaggregating our results by the different numeracy and literacy measures, significant differences in gender heterogeneity emerge (Table A4). First, consistent with our results on the aggregated literacy measure, males and females both benefit from the intervention for lower-order literacy skills, with smaller but insignificant differences in effect sizes for females. However, on the higher-order literacy skills, particularly on words read in the longer passage, we find that even though both males and females both benefit from the EdTech intervention, the difference in effects is significantly lower for females– treated females read on average 10 fewer words than their male counterparts. We find similar results on numeracy, with females with FD once again reporting significantly lower effect sizes on two higher-order numeracy skills–multiplication and division. We can therefore, conclude that our intervention has significantly higher effects for higher-order numeracy and literacy skills of males than of females with FD.

6.1.3 Heterogeneity by school grade (class)

We also examine whether there is any heterogeneity in treatment effects by the grade of the children with FD (Table 5). Examining this helps to shed light on the question whether children with FD at higher- or lower-level grades are more likely to benefit from the intervention (Table 5). We do not observe any significant differences in effects across school grades, suggesting that the technology intervention doesn't seem to be more beneficial for children at higher grades than it is for those at lower grades in either numeracy or literacy.

Table 5: Treatment heterogeneity by school grade

	Literacy	Numeracy
Treat x Grade 5	0.155	0.054
	(0.149)	(0.141)
Treat x Grade 4	0.064	-0.169
	(0.177)	(0.143)
Grade 5	0.002	0.085
	(0.126)	(0.101)
Grade 4	-0.135	0.087
	(0.136)	(0.099)
Treat	0.081	0.131
	(0.127)	(0.101)
Baseline score	0.607***	0.586***
	(0.040)	(0.039)
Clustered SEs (school)	Yes	Yes
Num. obs.	565	565
N Clusters	63	63

Notes: Table shows treatment effects by school grade. Both numeracy and literacy outcomes are standardized z-scores, which combines several observed measures from a standardized test. The literacy index includes: whether the child can read a simple phrase; number of correct questions related to the simple phrase (out of 2); the number of words the child can read from a 72-word passage; and number of correctly answered questions related to the longer passage (out of 5). The numeracy index comprises the number of correct responses to: four number identification questions; three questions on which number is bigger; four addition questions; three skip-pattern questions; four multiplication questions; and four division questions. Robust standard errors in parenthesis with the following significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Our disaggregated results (Table A5) provides some further insights to our findings. We find significantly larger positive effects on three of our our higher-order numeracy skills—addition, multiplication and division—for children in Grade 3 (represented by the co-efficient of "Treat"). We also observe that the differences in the magnitude of the effect is significantly larger for tow measures—the number of words read and number identification questions—for children in Grade 5 and significantly lower with regard to addition questions for children in Grade 4.

6.1.4 Heterogeneity by functional difficulty domain

We also explore FD heterogeneity to understand whether our treatment effects are different by functional difficulty domain, comparing each FD domain to the other (Table 6). For example, our Model 1 (Seeing- Lit), specifies the treatment effect comparing literacy outcomes for children with Seeing FD to the literacy outcomes of those with the other FDs combined. We do not find any significant marginal effects in either literacy or numeracy for any of our FD domains compared to the others.

Table 6: Treatment heterogeneity by functional difficulty (FD)

	Seeing		Hearing		Physical		Learning		Behavior	
	Lit	Num	Lit	Num	Lit	Num	Lit	Num	Lit	Num
Treat x Seeing FD	-0.121	-0.134								
	(0.183)	(0.156)								
Seeing FD	0.175	0.253*								
	(0.127)	(0.133)								
Treat x Hearing FD			-0.224	-0.028						
			(0.134)	(0.112)						
Hearing FD			0.253***	0.056						
			(0.082)	(0.076)						
Treat x Physical FD					0.384	0.114				
					(0.231)	(0.204)				
Physical FD					-0.429**	0.029				
					(0.163)	(0.147)				
Treat x Learning FD							-0.237	-0.077		
							(0.169)	(0.106)		
Learning FD							-0.000	-0.139*		
							(0.129)	(0.075)		
Treat x Behavior FD									0.341	0.084
									(0.241)	(0.245)
Behavior FD									-0.359*	-0.288
									(0.194)	(0.211)
Treat	0.175*	0.121	0.206**	0.109	0.126	0.091	0.297**	0.152**	0.130	0.099
	(0.088)	(0.076)	(0.086)	(0.071)	(0.081)	(0.067)	(0.117)	(0.075)	(0.080)	(0.060)
Baseline score	0.605***	0.588***	0.614***	0.604***	0.620***	0.607***	0.597***	0.578***	0.611***	0.598***
	(0.040)	(0.037)	(0.038)	(0.036)	(0.037)	(0.035)	(0.039)	(0.039)	(0.038)	(0.038)
Controls	No	No	No	No	No	No	No	No	No	No
Clustered SEs (school)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	564	564	564	564	564	564	564	564	564	564
N Clusters	63	63	63	63	63	63	63	63	63	63

Notes: Table shows treatment effect for a specific FD compared to other FDs. Both numeracy and literacy outcomes are standardized z-scores, which combines several observed measures from a standardized test. The literacy index includes: whether the child can read a simple phrase; number of correct questions related to the simple phrase (out of 2); the number of words the child can read from a 72-word passage; and number of correctly answered questions related to the longer passage (out of 5). The numeracy index comprises the number of correct responses to: four number identification questions; three questions on which number is bigger; four addition questions; three skip-pattern questions; four multiplication questions; and four division questions. We cluster our standard errors at the school level. Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

6.2 Possible mechanisms

6.2.1 Treatment effect for compliers

So far, our analysis has focused on the intention-to-treat (ITT) effects, that is, treatment assignment. One of the main mechanisms through which an EdTech intervention can be expected to influence learning outcomes is by interacting with the intervention itself. In our context, this would involve receiving the device and using it. We hypothesize a key expectation that children who receive the tablet (and app)–compliers– will exhibit significantly greater improvements in literacy outcomes post-intervention compared to those who do not receive the tablet (combining treatment group non-compliers and control group). To test this, we estimate the local average treatment effect (LATE) to isolate the impact of the intervention for compliers.

Table 7: Treatment effect for compliers

	Lit	Num
Received tablet	0.172^{**}	0.108^{*}
	(0.078)	(0.064)
Baseline score	0.610^{***}	0.611^{***}
	(0.040)	(0.035)
Controls	Yes	Yes
Clustered SEs	Yes	Yes
Num. obs.	565	565
N Clusters	63	63

Notes: Table shows outcomes by whether tablet was issued, and the number of days and months used. Both numeracy and literacy outcomes are standardized z-scores, which combines several observed measures from a standardized test. The literacy index includes: whether the child can read a simple phrase; number of correct questions related to the simple phrase (out of 2); the number of words the child can read from a 72-word passage; and number of correctly answered questions related to the longer passage (out of 5). The numeracy index comprises the number of correct responses to: four number identification questions; three questions on which number is bigger; four addition questions; three skip-pattern questions; four multiplication questions; and four division questions. We include gender, age, poverty likelihood, and female household head years of education as controls. Robust standard errors in parenthesis with the following significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

We identify two key findings. First, among compliers, the intervention has a significant impact on both literacy and numeracy outcomes. Second, the treatment effect is notably stronger among compliers. For literacy, the effect size in the ITT estimation is smaller and weaker at 0.15 standard deviations (90% confidence interval), whereas among compliers, it increases to 0.17 standard deviations with a stronger 95% confidence interval. Similarly, for numeracy, the ITT effect is small and non-significant (0.10 standard deviations, 85% confidence interval), but among compliers, the effect size is both larger and statistically significant at 0.11 standard deviations (90% confidence interval).

6.2.2 Changes in school perception and social interactions

An alternative mechanism through which EdTech interventions might influence literacy or numeracy rates is by altering students' interest in school or schooling activities, as well as their interactions with peers and teachers. However, as shown in Table 8, we do not observe any significant differences in interest in attending school, participation in school events, or conflicts with teachers or peers among the treatment group following the intervention.

Table 8: School participation and social interactions

	I like to...		I have conflcits with my...	
	go to school	help with activities at sch	friends in sch	teachers
Treat	-0.043	-0.036	-0.011	-0.003
	(0.035)	(0.046)	(0.021)	(0.019)
Baseline response	0.148*	-0.008	0.014	0.060
	(0.085)	(0.045)	(0.050)	(0.055)
Controls	No	No	No	No
Clustered SEs (school)	Yes	Yes	Yes	Yes
Num. obs.	565	565	565	565
N Clusters	63	63	63	63

Notes: Table shows child self-reported school participation and social interactions. Model 1 and 2 report changes in interest in attending school and participating in school events. Model 3 and 4 reports conflicts with friends and teachers, both at school. Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

6.2.3 Changes in time spent studying at home

Another potential mechanism through which the intervention might influence learning outcomes is by altering the amount of time children dedicate to studying. Changes in literacy or numeracy outcomes might, therefore, reflect increased effort rather than the direct impact of the technology itself. To examine this possibility, we analyze whether the intervention led to significant changes in the amount of time children spent studying at home. As part of our survey, children were asked to report the number of hours they spent studying or completing homework on the previous day. To account for potential variation in study hours by day of the week, we include day-of-week fixed effects in our estimation. The results, presented in Table 9, indicate no significant differences between the treatment and control groups in the amount of time children spent studying, either individually or with others in the household.

Table 9: Hours spent studying or doing homework

	Read by self	With others in HH
Treat	-0.026	0.043
	(0.091)	(0.040)
Baseline hours	0.147***	0.073*
	(0.050)	(0.041)
Controls	No	No
Clustered SEs	Yes	Yes
Day of week FEs	Yes	Yes
Num. obs.	564	565
N Clusters	63	63

Notes: Table shows child self-reported time spent (in hours) reading anything, writing or doing any homework while at home yesterday. In Model 1 and 2, we show time spent by child themselves, while in Models 3 and 4, we show time spent with others. Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

6.2.4 Changes in mental health and well-being

We also investigate whether the observed effects of the EdTech intervention for our learning outcomes outcomes might be attributed to changes in mental health or well-being arising from the provision of a sophisticated device. The rationale is that the endowment of such a device could potentially enhance participants' mental status, which, in turn, might influence learning outcomes. To assess mental health and well-being in our study, we used the Stirling Children's Well-being Scale (SCWBS), a validated instrument designed to evaluate emotional and mental well-being in educational and health contexts (Liddle and Carter, 2015). The SCWBS measures three core dimensions: positive emotional state, positive outlook, and social desirability, along with an aggregate metric that combines the first two dimensions.

As shown in Table 10, we find no significant changes in any of these measures, whether analyzed individually (Models 1–3) or as an aggregate (Model 4). These findings suggest that educational technology interventions do not appear to impact the mental health and well-being of the targeted children. Consequently, this provides additional evidence that changes in mental health status are unlikely to serve as a significant mechanism through which such interventions influence literacy or numeracy outcomes.

Table 10: Effects on mental heath and well-being

	Model 1	Model 2	Model 3	Model 4
	Positive emotions	Positive outlook	Social desirability	Aggregate Score
Treat	0.275	0.096	-0.107	0.369
	(0.429)	(0.457)	(0.234)	(0.782)
Baseline SCWBS score	0.036	0.051	0.162***	0.052
	(0.039)	(0.044)	(0.037)	(0.042)
Controls	No	No	No	No
Clustered SEs (school)	Yes	Yes	Yes	Yes
Num. obs.	565	565	565	565
N Clusters	63	63	63	63

Notes: Mental health measures according to the Stirling Children's Well-Being Scale (Liddle and Carter, 2015). The aggregate measure combines the first two dimensions—positive emotions and positive outlook—but excludes social desirability, as prescribed by the SCWBS. The analysis includes the child's baseline SCWBS score as a control variable but not any other covariates. Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

6.3 Sibling spillovers

Lastly, we investigate whether the intervention generated treatment spillovers to the siblings[9] of treated children. During the baseline and endline surveys, we administered our standardized test to 264 and 233 sibling respectively. However, given the inclusion criteria applied, there is minimal overlap in the siblings interviewed in both surveys—65 siblings in total. Although this final matched sample is relatively small, we demonstrate that it is balanced on baseline characteristics (Table A3), with the exception of a few numeracy metrics. These include identifying which number is bigger, addition, and skip pattern questions[10]. However, we are not overly concerned about these differences, as in all three cases, the siblings of the treated sample have lower baseline values. Furthermore, in our analysis, we include the baseline literacy or numeracy score to account for these differences and to improve the precision of our estimates.

We do not find evidence of significant spillover effects in numeracy or literacy outcomes among the siblings of treated children (Table 11). Our disaggregated results align with these findings (Table A6). For literacy, apart from a negative effect on the ability to read a simple phrase, there are no significant differences in post-intervention scores across the remaining four measures. Similarly, for numeracy, while we observe significant positive effects on addition, no significant effects are found for the five other numeracy measures.

[9]Details on the selection of siblings are discussed in Section 4.1.1

[10]The difference in skip pattern questions is only significant at the 90% confidence level.(Table A3)

Table 11: Sibling spillover effects

	Lit	Num
Treat	-0.100	0.199
	(0.212)	(0.134)
Baseline score	0.520***	0.573***
	(0.097)	(0.074)
Clustered SEs	Yes	Yes
Num. obs.	67	67
N Clusters	37	37

Notes: Table shows treatment effects by school grade. Both numeracy and literacy outcomes are standardized z-scores, which combines several observed measures from a standardized test. The literacy index includes: whether the child can read a simple phrase; number of correct questions related to the simple phrase (out of 2); the number of words the child can read from a 72-word passage; and number of correctly answered questions related to the longer passage (out of 5). The numeracy index comprises the number of correct responses to: four number identification questions; three questions on which number is bigger; four addition questions; three skip-pattern questions; four multiplication questions; and four division questions. Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

7 Conclusions

This study provides experimental evidence on the impact of an educational technology (EdTech) intervention on the literacy and numeracy outcomes of children with functional difficulties (FD) in a rural setting in Kenya. We implemented a cluster-randomized controlled trial in Homabay County, Kenya, with a sample of sixty five public primary schools. First, children in Grades 3–5 with functional difficulties in these schools were identified through a screening process by the County Disability Assessment Office (CDAO)–a government agency responsible for disability screening. Then, screened schools were randomized into treatment and control groups.

Our intervention involved distributing tablets preloaded with ANTON, a self-study educational app, to children in the treatment group. The tablets were accompanied by solar panels to support charging. The app offered curriculum-aligned content in multiple subjects and included gamification features to enhance engagement. Tablets were configured to restrict non-educational use and were distributed after training sessions for caregivers and children.

First, we find that the EdTech intervention significantly improves literacy outcomes, with effect sizes exceeding the median reported in similar studies. Specifically, treated children demonstrated a 0.15 standard deviation increase in literacy scores, driven primarily by improvements in lower-order literacy skills, such as the ability to read simple phrases and answer related comprehension questions. However, we find no significant effects for numeracy outcomes at the aggregate level, though in our disaggregated analysis we find positive impacts on higher-order numeracy skills, such as addition and multiplication.

Second, our analysis of heterogeneity highlights important variations in treatment effects across subgroups. Specifically on literacy, while both male and female students benefited from the intervention, the effect sizes were smaller for females– but statistically insignificant. Furthermore, females exhibited significantly lower gains in higher-order literacy and numeracy skills compared to males. These findings suggest that while EdTech interventions hold promise for improving learning outcomes, gender-specific barriers may limit their effectiveness for female students with FD. Noteworthy, we found no significant differences in treatment effects across school grades or functional difficulty domains, indicating that the intervention was broadly beneficial across these dimensions.

Third, our exploration of potential mechanisms highlights the importance of compliance with the intervention. Children who actively used the tablets (compliers) exhibited stronger treatment effects, with significant improvements in both literacy and numeracy outcomes. This suggests that the effectiveness of EdTech interventions is closely tied to the extent of device usage. However, we found no evidence that the intervention influenced other potential mechanisms, such as changes in study time, school perceptions, social interactions, or mental health and well-being. This reinforces the conclusion that the observed improvements in learning outcomes are directly attributable to the use of the educational technology rather than indirect behavioral or psychological changes. Finally, we do not find any significant spillover effects to the sibling of our treated sample.

Our results have important implications for policymakers and practitioners seeking to address the learning gaps faced by children with disabilities in low-income settings. The positive effects of the EdTech intervention on literacy outcomes demonstrate the potential of technology-enabled learning tools to complement traditional educational approaches, particularly for children with functional difficulties who may face barriers to regular school attendance or require personalized learning support. To maximize the impact of EdTech interventions, policymakers should prioritize strategies that promote consistent device usage.

References

Agrawal, K., Athey, S., Kanodia, A., & Palikot, E. (2022, December 4). Personalized recommendations in EdTech: Evidence from a randomized controlled trial. Retrieved January 27, 2023, from http://arxiv.org/abs/2208.13940

Angrist, N., Bergman, P., & Matsheng, M. (2022). Experimental evidence on learning using low-tech when school is out [Publisher: Nature Publishing Group]. *Nature Human Behaviour*, *6*(7), 941–950. https://doi.org/10.1038/s41562-022-01381-z

Banergee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., & Walton, M. (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations of "teaching at the right level" in india. http://www.nber.org/papers/w22746

Bastawrous, A., & Armstrong, M. J. (2013). Mobile health use in low- and high-income countries: An overview of the peer-reviewed literature [Publisher: SAGE Publications]. *Journal of the Royal Society of Medicine*, *106*(4), 130–142. https://doi.org/10.1177/0141076812472620

Büchel, K., Jakob, M., Kühnhanss, C., Steffen, D., & Brunetti, A. (2022). The relative effectiveness of teachers and learning software: Evidence from a field experiment in el salvador [Publisher: The University of Chicago Press]. *Journal of Labor Economics*, *40*(3), 737–777. https://doi.org/10.1086/717727

Escueta, M., Nickow, A. J., Oreopoulos, P., & Quan, V. (2020). Upgrading education with technology: Insights from experimental research. *Journal of Economic Literature*, *58*(4), 897–996. https://doi.org/10.1257/jel.20191507

Evans, D. K., & Yuan, F. (2022). How big are effect sizes in international education studies? [Publisher: American Educational Research Association]. *Educational Evaluation and Policy Analysis*, *44*(3), 532–540. https://doi.org/10.3102/01623737221079646

Glewwe, P., & Kremer, M. (2006, January 1). Chapter 16 schools, teachers, and education outcomes in developing countries. In E. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (pp. 945–1017). Elsevier. https://doi.org/10.1016/S1574-0692(06)02016-2

Glewwe, P., Krutikova, S., & Rolleston, C. (2017). Do schools reinforce or reduce learning gaps between advantaged and disadvantaged students? evidence from vietnam and peru [Publisher: The University of Chicago Press]. *Economic Development and Cultural Change*, *65*(4), 699–739. https://doi.org/10.1086/691993

Glewwe, P., Park, A., & Zhao, M. (2016). A better vision for development: Eyeglasses and academic performance in rural primary schools in china. *Journal of Development Economics*, *122*, 170–182. https://doi.org/10.1016/j.jdeveco.2016.05.007

Government of Kenya. (2019). *Basic education statistical booklet, 2019*. Retrieved November 18, 2022, from https://africacheck.org/sites/default/files/Kenya-Basic-Education-Statistical-Booklet-2019.pdf

Grimm, M., & Hartwig, R. (2018, November 19). Unblurring the market for vision correction: A willingness to pay experiment in rural burkina faso. Retrieved October 28, 2022, from https://papers.ssrn.com/abstract=3286189

Hanafin, J., Shevlin, M., Kenny, M., & Neela, E. M. (2007). Including young people with disabilities: Assessment challenges in higher education. *Higher Education*, *54*(3), 435–448. https://doi.org/10.1007/s10734-006-9005-9

Harris, M., & Terlektsi, E. (2011). Reading and spelling abilities of deaf adolescents with cochlear implants and hearing aids. *The Journal of Deaf Studies and Deaf Education*, *16*(1), 24–34. https://doi.org/10.1093/deafed/enq031

Kenya National Bureau of Statistics. (2020). *2019 kenya population and housing census volume IV: Distribution of population by socio-economic characteristics* [Kenya national bureau of statistics]. Retrieved August 3, 2023, from https://www.knbs.or.ke/download/2019-kenya-population-and-housing-census-volume-iv-distribution-of-population-by-socio-economic-characteristics/

Lamichhane, K. (2013). Disability and barriers to education: Evidence from nepal [Number: 4 Publisher: Stockholm University Press]. *Scandinavian Journal of Disability Research*, *15*(4), 311–324. https://doi.org/10.1080/15017419.2012.703969

Liddle, I., & Carter, G. F. (2015). Emotional and psychological well-being in children: The development and validation of the stirling children's well-being scale [Publisher: Routledge _eprint: https://doi.org/10.1080/02667363.2015.1008409]. *Educational Psychology in Practice*, *31*(2), 174–185. https://doi.org/10.1080/02667363.2015.1008409

Lockwood, F. (2006). Hearing impaired children in the mainstream classroom: The effects on literacy and learning. In *Differentiation and diversity in the primary school* (pp. 107–116). Routledge.

Ma, X., Zhou, Z., Yi, H., Pang, X., Shi, Y., Chen, Q., Meltzer, M. E., Cessie, S. l., He, M., Rozelle, S., Liu, Y., & Congdon, N. (2014). Effect of providing free glasses on children's educational outcomes in china: Cluster randomized controlled trial [Publisher: British Medical Journal Publishing Group Section: Research]. *BMJ*, *349*, g5740. https://doi.org/10.1136/bmj.g5740

McKinney, E. L., & Swartz, L. (2016). Life in special schools in south africa: Voices of former students. *International Journal of Disability, Development and Education*, *63*(3), 309–321. https://doi.org/10.1080/1034912X.2015.1089980

Moriña Díez, A. (2010). School memories of young people with disabilities: An analysis of barriers and aids to inclusion [Publisher: Routledge]. *Disability & Society*, *25*(2), 163–175. https://doi.org/10.1080/09687590903534346

Odhiambo, F., & Gunther, I. (2023). Does edTech use improve numeracy and literacy skills and mental well-being for children with functional difficulties? evidence from kenya [Institution: American Economic Association]. https://doi.org/10.1257/rct.12044-1.0

Piper, B., Zuilkowski, S. S., Kwayumba, D., & Strigel, C. (2016). Does technology improve reading outcomes? comparing the effectiveness and cost-effectiveness of ICT interventions for early grade reading in kenya. *International Journal of Educational Development*, *49*, 204–214. https://doi.org/10.1016/j.ijedudev.2016.03.006

Rodriguez-Segura, D. (2022). EdTech in developing countries: A review of the evidence. *The World Bank Research Observer*, *37*(2), 171–203. https://doi.org/10.1093/wbro/lkab011

Trani, J.-F., Bakhshi, P., & Nandipati, A. (2012). 'delivering' education; maintaining inequality. the case of children with disabilities in afghanistan [Publisher: Routledge _eprint: https://doi.org/10.1080/0305764X.2012.706259]. *Cambridge Journal of Education*, *42*(3), 345–365. https://doi.org/10.1080/0305764X.2012.706259

Wang, X., Congdon, N., Ma, Y., Hu, M., Zhou, Y., Liao, W., Jin, L., Xiao, B., Wu, X., Ni, M., Yi, H., Huang, Y., Varga, B., Zhang, H., Cun, Y., Li, X., Yang, L., Liang, C., Huang, W., … Ma, X. (2017). Cluster-randomized controlled trial of the effects of free glasses on purchase of children's glasses in china: The PRICE (potentiating rural investment in children's eyecare) study [Publisher: Public Library of Science]. *PLOS ONE*, *12*(11), e0187808. https://doi.org/10.1371/journal.pone.0187808

World Bank. (2018). *World development report 2018: Learning to realize education's promise.* https://doi.org/10.1596/978-1-4648-1096-1

Yuwono, I., & Okech, J. B. (2021). The classroom impact of trained special needs education teachers in selected schools: An evaluation study. *Frontiers in Education*, *6*. Retrieved October 31, 2022, from https://www.frontiersin.org/articles/10.3389/feduc.2021.630806

Zhang, H., & Holden, S. T. (2023). Numeracy skills learning of children in africa:—are disabled children lagging behind? [Publisher: Public Library of Science]. *PLOS ONE*, *18*(4), e0284821. https://doi.org/10.1371/journal.pone.0284821

A Appendices

A.1 Intervention

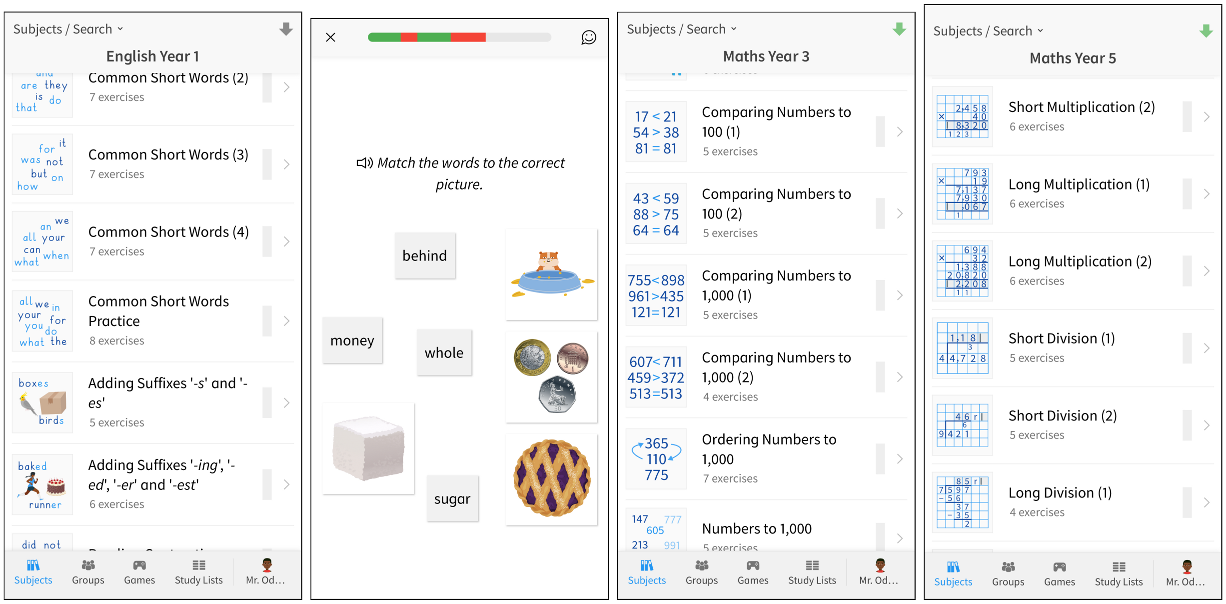
Figure A1: Tablet

Notes: Tablet issued to study participants for studying the educational material.

Figure A2: Solar panel

Notes: Solar panel issued to study participants for charging the tablets.

Figure A3: Anton App



Notes: Screenshots of some learning content available on the ANTON app. Image 1 displays English Year 1 materials, while Image 2 features "Common Exception Words" exercises from English Year 2. Image 3 showcases Math Year 3 exercises, and Image 4 highlights multiplication and division exercises from Math Year 5. The authors do not own or claim rights to the content shown or the app itself.

A.2 Attrition

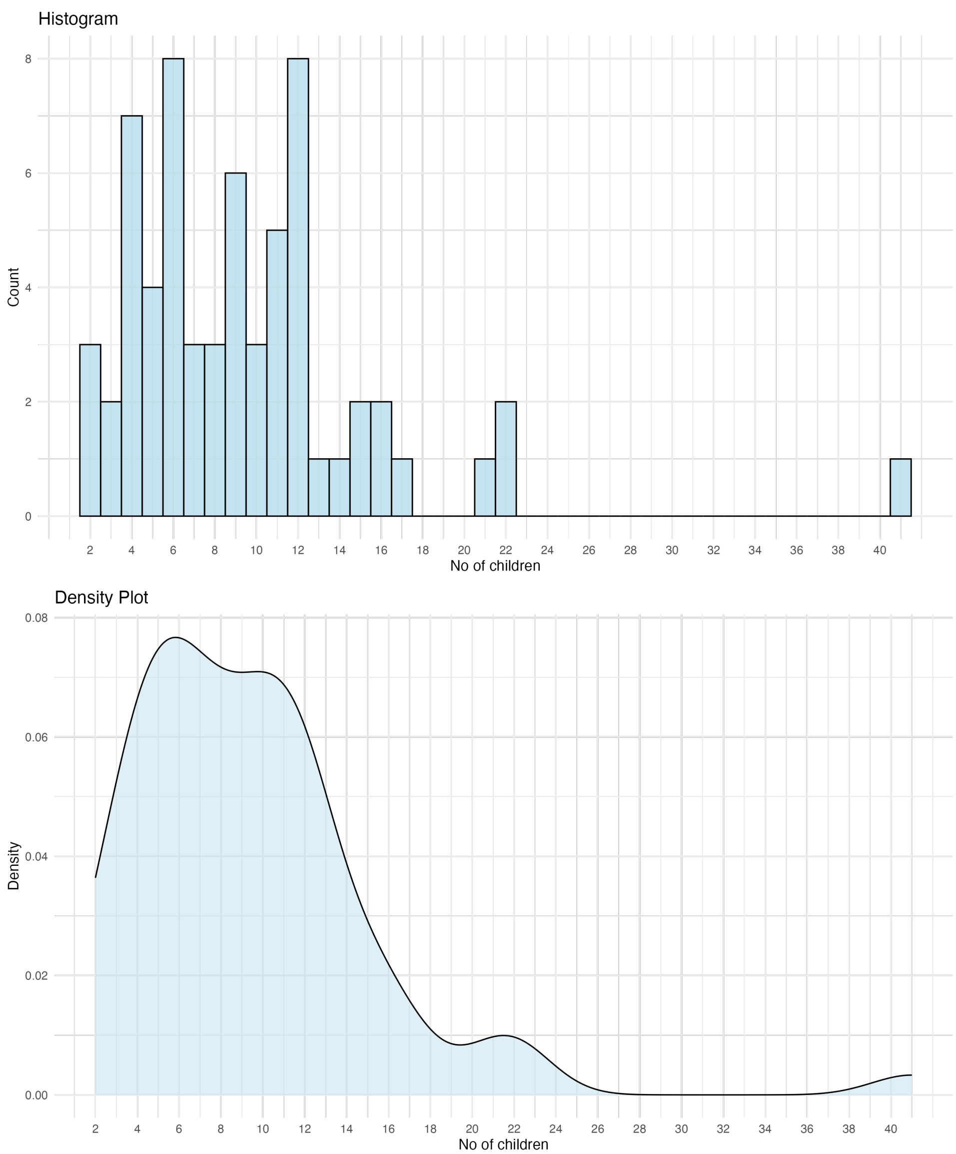
Table A1: Attrition Table

	Control	Treated	p-value
n	306	318	
Parent survey (mean (SD))	0.06 (0.24)	0.05 (0.22)	0.505
Child survey (mean (SD))	0.05 (0.22)	0.03 (0.18)	0.232

Notes: Table shows attrition at endline, reporting both parent survey and child survey attrition. Attrition rates are lower for the child survey given in some cases the parent/caregiver provided consent for the survey to be administered on the child via phone, even though the parent/caregiver themselves were absent for an extended period.

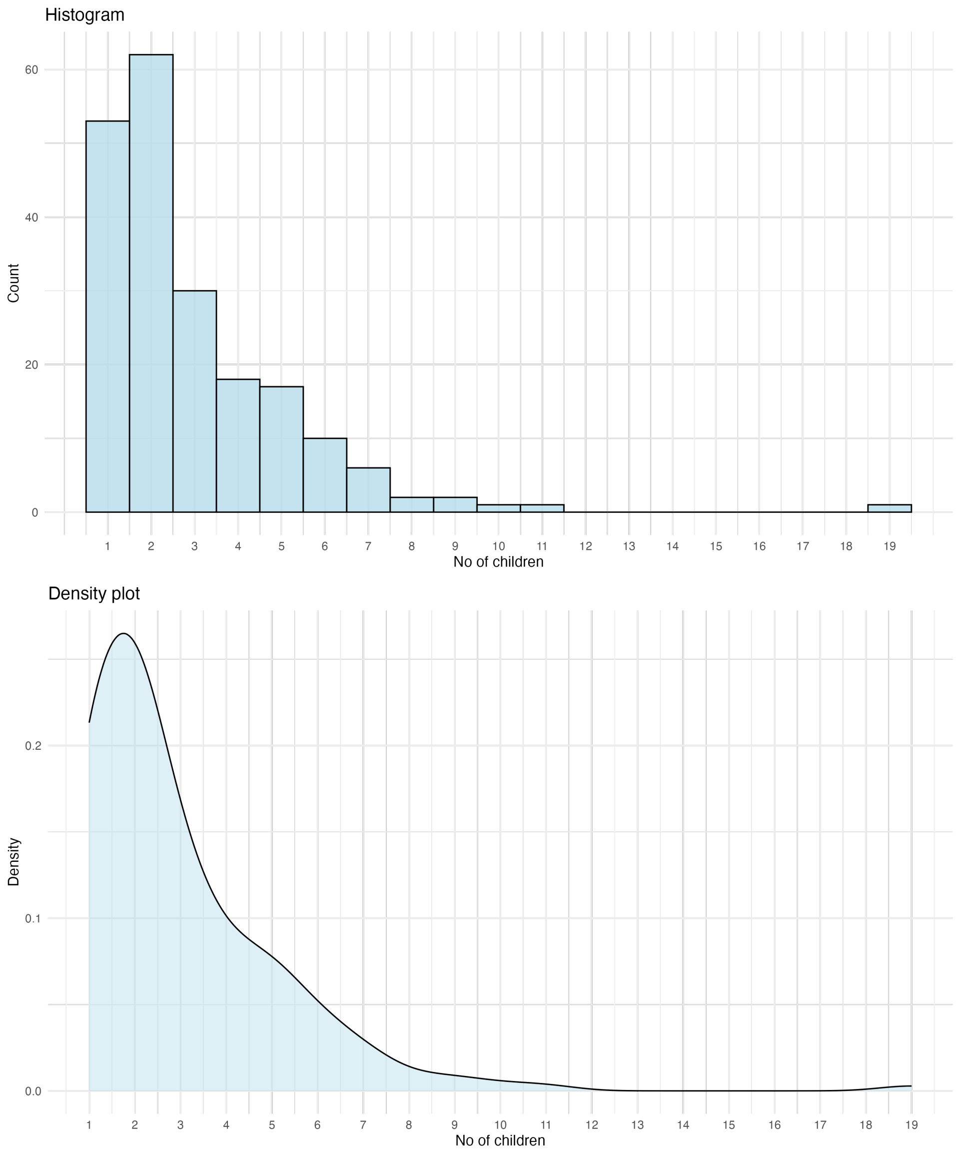
A.3 Distribution of study children by class and school

Figure A4: Distribution by school



Notes: The figure plots histogram and density plot to show the distribution of children per school

Figure A5: Distribution by class



Notes: The figure plots histogram and density plot to show the distribution of children per class in a school. In each school, we sample children from Grades 3, 4 and 5.

A.4 Functional difficulty (FD) screening

Table A2: FD Assessment Status by Class

Child Grade	With FD	Without FD	Referred	Total
Class 1	4.3% (36)	9.5% (4)	4.0% (3)	4.5% (43)
Class 2	7.6% (63)	7.1% (3)	9.3% (7)	7.7% (73)
Class 3	27.0% (224)	19.0% (8)	24.0% (18)	26.4% (250)
Class 4	22.3% (185)	16.7% (7)	6.7% (5)	20.8% (197)
Class 5	22.4% (186)	14.3% (6)	26.7% (20)	22.4% (212)
Class 6	6.7% (56)	21.4% (9)	9.3% (7)	7.6% (72)
Other	9.7% (81)	11.9% (5)	20.0% (15)	10.7% (101)
Total	100.0% (831)	100.0% (42)	100.0% (75)	100.0% (948)

Notes: Table shows a summary of the children who were screened by the CDAOs and the final Functional Disability (FD) status by Grade. Only children in Grades 3,4 and 5 were targeted for screening, hence the higher proportions in those Grades. Some caregivers of children in the other Grades also showed up for the assessment.

A.5 Sibling balance table

Table A3: Balance Table (sibling interviewed at both baseline and endline)

Variable	Control	Treated	p-value
Background			
Female	0.52	0.54	0.52
Age	11.67	11.32	0.13
Poverty likelihood	0.67	0.67	0.53
Schooling			
Grade (school level)	4.24	4.20	0.84
Days attended school last week	3.86	4.07	0.08
Hours studying at home yesterday	0.84	0.93	0.14
Literacy			
Can read simple phrase	0.70	0.64	0.28
Number of words read correctly (out of 72)	41.31	34.69	0.11
No. of comprehension questions correct (out of 5)	2.25	2.14	0.68
Numeracy			
Number identification (out of 4)	2.90	2.72	0.21
Which number is bigger (out of 3)	2.55	2.31	0.02
Addition (out of 4)	2.44	2.04	0.03
Skip pattern (out of 3)	1.69	1.46	0.09
Multiplication (out of 4)	1.78	1.70	0.68
Division (out of 3)	1.29	1.24	0.77
N	28.00	39.00	NA

Notes: Table shows the balance table of the siblings of our study children at baseline, and only includes siblings who could be matched at endline. The specific literacy and numeracy test questions administered to the participants are shown in Appendix A.8 and Appendix A.9 respectively.

A.6 Treatment heterogeneity

Table A4: Treatment heterogeneity by gender- disaggregated

	Literacy					Numeracy					
	Read	Q1	Q2	Words	Q3-Q7	Num	Big	Add	Skip	Mult	Div
Treat x Female	-0.029	-0.030	-0.045	-9.854*	-0.172	-0.054	0.000	-0.073	-0.248	-0.592**	-0.502*
	(0.064)	(0.071)	(0.068)	(5.210)	(0.341)	(0.170)	(0.131)	(0.214)	(0.251)	(0.245)	(0.255)
Treat	0.092*	0.110**	0.097*	8.860**	0.341	0.094	-0.013	0.256	0.216	0.577***	0.465**
	(0.050)	(0.049)	(0.050)	(3.677)	(0.241)	(0.154)	(0.108)	(0.165)	(0.163)	(0.159)	(0.181)
Female	0.101**	0.091*	0.145***	12.754***	0.468*	0.027	0.060	0.065	-0.062	0.195	0.201
	(0.041)	(0.054)	(0.048)	(3.775)	(0.271)	(0.136)	(0.106)	(0.165)	(0.189)	(0.176)	(0.197)
Baseline score	0.498***	0.470***	0.467***	0.566***	0.624***	0.472***	0.410***	0.475***	0.684***	0.488***	0.551***
	(0.046)	(0.039)	(0.046)	(0.042)	(0.040)	(0.048)	(0.050)	(0.038)	(0.047)	(0.042)	(0.048)
Controls	No	No	No	No	No	No	No	No	No	No	No
Clustered SEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	565	565	565	565	565	565	565	565	565	565	565
N Clusters	63	63	63	63	63	63	63	63	63	63	63

Notes: Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01. The literacy measures include: a binary indicator for whether the child can read a simple phrase; whether they correctly answer two comprehension questions related to the simple phrase; the number of words the child can read from a 72-word passage; and a binary indicator for whether they correctly answer five comprehension questions related to the longer passage. The numeracy measures comprises: the number of correct responses to four number identification questions; the number of correct responses to three questions on which number is bigger; the number of addition questions (out of four) answered correctly; the number of correct responses to three skip-pattern questions; the number of correct answers to four multiplication questions; and the number of correct responses to three division questions. We cluster our standard errors at the school level.

Table A5: Treatment heterogeneity by school grade- disaggregated

	Literacy					Numeracy					
	Read	Q1	Q2	Words	Q3-Q7	Num	Big	Add	Skip	Mult	Div
Treat x Grade 5	0.036	0.023	0.034	10.700**	0.191	0.405*	0.064	-0.017	0.175	-0.079	-0.119
	(0.076)	(0.085)	(0.079)	(4.764)	(0.357)	(0.239)	(0.181)	(0.297)	(0.240)	(0.294)	(0.277)
Treat x Grade 4	0.041	-0.007	-0.013	7.481	-0.062	0.076	-0.246	-0.543**	-0.203	-0.277	-0.488
	(0.083)	(0.092)	(0.093)	(5.715)	(0.413)	(0.235)	(0.170)	(0.266)	(0.270)	(0.312)	(0.311)
Grade 5	0.026	0.065	0.030	-3.041	0.176	-0.021	0.196	0.311	0.254	0.403**	0.428**
	(0.062)	(0.068)	(0.064)	(3.976)	(0.287)	(0.193)	(0.159)	(0.201)	(0.172)	(0.199)	(0.190)
Grade 4	-0.056	-0.016	-0.062	-7.407	-0.118	0.040	0.075	0.445**	0.274	0.144	0.334
	(0.066)	(0.073)	(0.070)	(4.573)	(0.266)	(0.176)	(0.136)	(0.176)	(0.188)	(0.223)	(0.223)
Treat	0.052	0.087	0.067	-1.199	0.210	-0.084	0.034	0.387*	0.115	0.437**	0.438**
	(0.063)	(0.064)	(0.063)	(4.321)	(0.308)	(0.164)	(0.133)	(0.206)	(0.177)	(0.206)	(0.204)
Baseline score	0.501***	0.464***	0.467***	0.569***	0.614***	0.450***	0.391***	0.446***	0.639***	0.453***	0.506***
	(0.046)	(0.042)	(0.049)	(0.041)	(0.043)	(0.048)	(0.051)	(0.040)	(0.052)	(0.044)	(0.050)
Gr3 T_c End Mean	0.585	0.500	0.538	33.915	1.377	2.792	2.226	1.764	1.491	1.151	0.840
Controls	No	No	No	No	No	No	No	No	No	No	No
Clustered SEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	565	565	565	565	565	565	565	565	565	565	565
N Clusters	63	63	63	63	63	63	63	63	63	63	63

Notes: Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01. The literacy measures include: a binary indicator for whether the child can read a simple phrase; whether they correctly answer two comprehension questions related to the simple phrase; the number of words the child can read from a 72-word passage; and a binary indicator for whether they correctly answer five comprehension questions related to the longer passage. The numeracy measures comprises: the number of correct responses to four number identification questions; the number of correct responses to three questions on which number is bigger; the number of addition questions (out of four) answered correctly; the number of correct responses to three skip-pattern questions; the number of correct answers to four multiplication questions; and the number of correct responses to three division questions. We cluster our standard errors at the school level.

Table A6: Sibling spillover effects, disaggregated

	Literacy					Numeracy					
	Read	Q1	Q2	Words	Q3-Q7	Num	Big	Add	Skip	Mult	Div
Treat	-0.158*	-0.039	-0.005	-0.595	-0.296	-0.057	0.308	0.285	0.576*	0.118	0.198
	(0.090)	(0.102)	(0.117)	(7.911)	(0.535)	(0.237)	(0.191)	(0.217)	(0.308)	(0.316)	(0.263)
Baseline score	0.296***	0.406***	0.385***	0.479***	0.499***	0.441***	0.168	0.638***	0.744***	0.444***	0.405***
	(0.107)	(0.121)	(0.114)	(0.093)	(0.112)	(0.103)	(0.139)	(0.093)	(0.145)	(0.095)	(0.096)
Sibling T_c End Mean	0.618	0.545	0.500	34.782	2.200	2.645	1.855	1.973	1.673	1.518	1.173
Clustered SEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	67	67	67	67	67	67	67	67	67	67	67
N Clusters	37	37	37	37	37	37	37	37	37	37	37

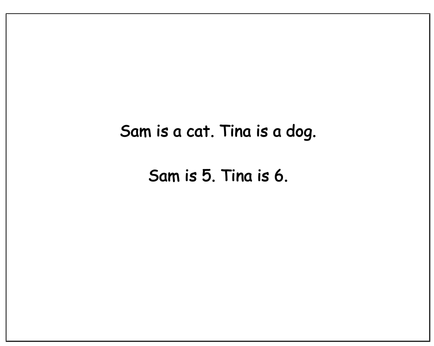
Notes: Table shows the spillover effects to siblings of our treated sample. Only baseline siblings who could be matched at endline are included. Table A3 presents their baseline characteristics and balance across treatment groups. The literacy index includes: whether the child can read a simple phrase; number of correct questions related to the simple phrase (out of 2); the number of words the child can read from a 72-word passage; and number of correctly answered questions related to the longer passage (out of 5). The numeracy index comprises the number of correct responses to: four number identification questions; three questions on which number is bigger; four addition questions; three skip-pattern questions; four multiplication questions; and four division questions. We cluster our standard errors at the school level (of the study child). Robust standard errors in parenthesis with the following significance levels: *p<0.1; **p<0.05; ***p<0.01.

A.7 Sibling spillovers

A.8 Literacy tests

A.8.1 Short passage

Figure A6: Short passage



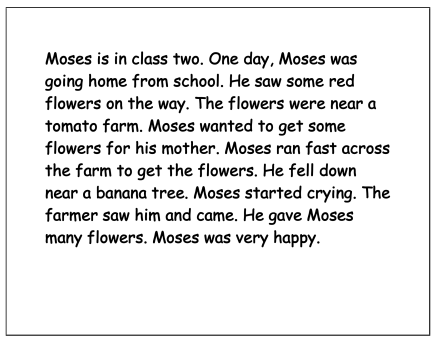
Notes: Short passage displayed to the participants as placards. Adapted from the Multiple Indicator Cluster Survey (MICS7)

Table A7: Short passage questions

Question		Response
1. How old is Sam?	0	Other answer
	1	5 years old
	998	Did not answer or No response
2. Who is older: Sam or Tina?	0	Sam
	1	Tina
	998	Did not answer or No response

A.8.2 Long passage

Figure A7: Long passage



Moses is in class two. One day, Moses was going home from school. He saw some red flowers on the way. The flowers were near a tomato farm. Moses wanted to get some flowers for his mother. Moses ran fast across the farm to get the flowers. He fell down near a banana tree. Moses started crying. The farmer saw him and came. He gave Moses many flowers. Moses was very happy.

Notes: Long passage displayed to the participants as placards. Adapted from the Multiple Indicator Cluster Survey (MICS7)

Table A8: Long passage questions

Question		Response
3. What class is Moses in?	0	Other response
	1	Two
	998	Did not answer or No response
4. What did Moses see on the way home?	0	Other response
	1	Flowers
	998	Did not answer/ No response
5. Why did Moses start crying?	0	Other response
	1	Because he fell
	998	Did not answer or No response
6. Where did Moses fall?	0	Other response
	1	Near a banana tree
	998	Did not answer or No response
7. Why was Moses happy?	0	Other response
	1	Because the farmer gave him many flowers or because he had flowers to give to his mother
	998	Did not answer or No response

A.9 Numeracy tests

Figure A8: Number identification questions

9
48
216
1,125

Notes: Child asked to point each number and tell what the number is. Adapted from the Multiple Indicator Cluster Survey (MICS7)

Figure A9: Which number is bigger questions

7 5	**9 65**	**159 175**

Notes: Child asked to identify which number is bigger. Adapted from the Multiple Indicator Cluster Survey (MICS7)

Figure A10: Addition questions

3 + 2 =	**34 + 8 =**	**17 + 24 =**	**115 + 230 =**

Notes: Placards with set of addition questions. Adapted from the Multiple Indicator Cluster Survey (MICS7)

Figure A11: Skip pattern questions

1 2 __ 4	3 6 __ 12	5 8 11 __	44 55 66 __

Notes: Child asked to identify which number is missing from the sequence of numbers. Adapted from the Multiple Indicator Cluster Survey (MICS7). Skip pattern question 3, 6, _, 12, is only asked at endline.

Figure A12: Multiplication questions

1 x 2 =	2 x 3 =	4 x 8 =	12 x 4 =

Notes: Placards with set of multiplication questions. Adapted from the Multiple Indicator Cluster Survey (MICS7)

Figure A13: Division questions

4 ÷ 1 =	6 ÷ 2 =	16 ÷ 4 =	225 ÷ 15 =

Notes: Placards with set of division questions. Adapted from the Multiple Indicator Cluster Survey (MICS7). Division question 225 divided by 15 is only asked at endline.