Null Impacts of Education and Information Interventions on Children's Cognitive and Noncognitive Skills: Experimental Evidence from Poor Urban Households in the Philippines

Andrew S. Griffen* Yukichi Mano† Makiko Nakamuro‡
Takaaki Okumura§ Yasuyuki Sawada¶

Abstract

We use a randomized-control trial (RCT) to assess the impact on children's cognitive and noncognitive skills of several educational interventions (drawing, storytelling, and e-learning) combined with a parental information intervention about the importance of educational investments. We found no impacts across a range of outcomes for either the educational intervention or the educational plus parental information intervention. We discuss potential reasons for the null result and directions for future research and policy.

Introduction

Poor parents and guardians in developing countries often underinvest in their children due to a lack of financial resources. However, incomplete credit markets, riskiness of such investments, irreversibility, and limited information about returns can also all lower potentially efficient investments in human capital (Becker et al. 2018). In principle, governments in developing countries can address some of these market failures however one of their main policy levers (schools) also faces its own set of challenges in effectively producing human capital (Kremer, Brannen, and Glennerster 2013; Glewwe and Muralidharan 2016; Snilstveit et al. 2015; Evans and Popova 2016; Glewwe et al. 2021). Financial constraints, infrastructure limitations, large enrollments, and large learner-teacher ratios all make teaching and remediation difficult[1].

Learners also frequently fall behind early in school which can affect the entire profile of investment decisions by children, parents, and schools (Cunha et al. 2006; Cunha and Heckman 2007; Cunha, Heckman, and Schennach 2010). This can rationalize dropout decisions and a lack of motivation as well as lead to issues with remediation and negative spillovers within the classroom. Poverty is also extremely challenging for mental health

*Faculty of Economics, University of Tokyo, griffen@e.u-tokyo.ac.jp
†Graduate School of Economics, Hitotsubashi University, yukichi.mano@r.hit-u.ac.jp
‡Faculty of Policy Management, Keio University, makikon@sfc.keio.ac.jp
§Nippon Sports Science University, okumura@nittai.ac.jp
¶Faculty of Economics, University of Tokyo, sawada@e.u-tokyo.ac.jp

[1]The two elementary schools in our study have such large enrollments that they split the day in half with some learners attending during a morning session typically from 5:40 - 12:00 and other learners attending an afternoon session typically from 12:00 to 18:20 (these times vary slightly by grade level). In addition, class sizes are often as high 60 learners per teacher. In the Philippines, they use the term *learner* in place of student and we adhere to their preferred nomenclature in this paper.

(Haushofer and Fehr 2014; Ridley et al. 2020; Barker et al. 2022; Jin et al. 2024) and economists have been increasingly recognizing the importance of cognitive and noncognitive development not only for economic outcomes but also for the production of skills themselves (Health Organization" 1999; Heckman, Stixrud, and Urzua 2006; Almlund et al. 2011). Finally, urban poverty in developing countries is often quite concentrated, which can create strong neighborhood and peer effects that present both a challenge and an opportunity for policymakers (Marx, Stoker, and Suri 2013). Together these issues can exacerbate to create a perfect storm for inequality and the intergenerational persistence of poverty.

To understand some potential policy approaches to these issues in a poor urban environment, we investigated whether supplementary educational interventions interacted with a parental information intervention might change the trajectory of children's skill development in a poor urban community located in Rizal Province in Manila, Philippines. In our study we worked with two elementary schools and one high school. Children were randomly assigned to one of three groups: educational intervention only, educational plus parental information intervention, or control[2]. The two treated groups received the same one hour per week educational intervention of either drawing, storytelling, or e-learning that depended on the learner's grade level[3]. The parental intervention provided an initial pre-intervention information session (with periodic follow-ups) on children's skill development, the returns to schooling, and the importance of savings and investment for education. Given the limited resources of both schools and families in the context, we tried to design these interventions to be feasible but also sustainable.

Disappointingly we found no impacts of our intervention across a range of human capital outcomes for children. This includes even from several years of follow-up data. We hope this paper can be useful to the literature in avoiding the *file drawer problem* (Andrews and Kasy 2019; Abadie 2020; Chopra et al. 2024) and to give a more complete picture of when and why such interventions do *not* appear to be effective. Particularly when viewing development as an engineering problem null results from RCTs are quite instructive (Karlan and Appel 2017).

Our findings are in contrast to the literature which typically finds positive impacts of e-learning interventions on mathematics test scores in developing countries[4]. Information about the returns to education have been demonstrated to produce positive impacts in some contexts (Nguyen 2008; Jensen 2010)[5]. RCTs for storytelling and drawing interventions appear to be completely novel in the literature[6]. E-learning has been much more studied but generalization and context of its effectiveness are not well understood (Kremer, Brannen, and Glennerster 2013; Bulman and Fairlie 2016). We are also not aware

[2]The proposed RCT was reviewed and approved internally at the University of Tokyo by an ethics review committee on 14-5-2015 (examination number 15-19) prior to the data collection. Documentation can be provided upon request.

[3]These type of interventions are variously called e-learning, computer aided instruction (CAI), and computer aided learning (CAL) in the literature (Bulman and Fairlie 2016).

[4]Programs that simply give technology such as computers or tablets tend not to find any impacts (Barrera-Osorio and Linden 2009) whereas e-learning software seems capable of producing large impacts (Banerjee et al. 2007; Linden 2008; Lai et al. 2013, 2015; Mo, Zhang, Wang, et al. 2014; Mo, Zhang, Luo, et al. 2014; Mo et al. 2015; Bai et al. 2016; Muralidharan, Singh, and Ganimian 2019)

[5]Instead of a very general information treatment, more papers seem to examine individualized information treatment such as providing parents clearer information about their child's performance in school. Interestingly it seems that lower income parents are less responsive to such information interventions (Boneva and Rauh 2018; Dizon-Ross 2019)

[6]There are claims that these types of arts interventions will provide spillovers to other types of skill development (Simpson Steele 2016). Despite the evidence base for such claims being fairly weak (Goldstein, Vincent-Lancrin, and Winner 2013), policies, including those in Japan, are predicated on exactly such mechanisms. The storytelling RCTs we found in the literature are more oriented around promoting health behaviors of adults and teenagers, which is qualitatively quite different than our intervention.

of research examining interactions between parental information and child educational intervention. We provide more details below when discussing the inventions.

We also try to draw some specific conclusions and lessons from our experience conducting the intervention. A positive side effect of our project is that we continue to work with local school district officials who are very interested in conducting impact analyses on different programs within their district and with a local NGO called the Salt Payatas Foundation Philippines (*Salt Payatas*) who continue their efforts to improve the lives of poor children in the Philippines. We hope this collaborative work will yield further capacity building and policy insights that can be helpful to improve the lives of children in the Philippines.

Background and context

Our intervention occurred in Kasiglahan Village (KV), which is located in Rizal Province in metro Manila. Although initially the village was a resettlement area for victims of the Payatas dumpsite landslide, the area has subsequently grown significantly[7]. Our intervention was designed based on consultation with school district officials in the Department of Education in Rizal Province and with the Salt Payatas who have been working with women and children in Payatas and KV since 1995[8].

Preliminary interviews with parents, teachers, volunteers, and the children themselves provided background information about the challenges that the children faced in their lives and helped to design the intervention. Our observations were as follows. First, some children have already stopped attending school as early as G1 and G2[9]. Second, many learners do not show much interest in reading or in school more generally. Even at the beginning of their school careers, lack of motivation and interest were already salient. Third, some learners lacked basic skills such as ability to spell their own name or to perform simple arithmetic calculations. These literacy and numeracy issues obviously compound and make learning in higher grades difficult. Furthermore, given the extremely large classes of approximately 60 learners per class, it is often difficult for teachers to provide individualized remediation or to motivate those who are not interested in school. Trying to increase motivation and the use of technology for large-scale remediation were identified early on as potential areas for intervention.

Given these facts on the ground, we worked with subject-matter experts to design three types of educational interventions: drawing, story telling, and e-learning. The assigned intervention depended on the grade of the child at the time of random assignment. Drawing was for learners initially in G1-G2, storytelling for those in G3-G4, and e-learning for G5 and above. Learners would continue to receive the same type of intervention for the duration of the study. Most of the interventions were for one hour per week, were extracurricular, and were around lunch time (either after the morning school session or

[7] On July 10, 2000 there was a devastating landslide of a garbage dump that had been used for scavenging by squatters living in a nearby area called Payatas. Although officially 218 people died, the total was perhaps much higher and many people's homes were either destroyed or declared at-risk and uninhabitable. A national outcry followed and the efforts of activists and children affected by the tragedy such as the Bangkang papel boys led to resettlement efforts.

[8] Salt Payatas provides educational scholarships to children and empowers women through production of craft goods for sale. These activities are partly financed by offering study tours of local areas. However, Salt Payatas felt that these scholarships were often insufficient to stop children from dropping out or performing well in school. This led to efforts to find other solutions and members of Salt Payatas made contact with researchers in Japan to begin thinking about what studies or interventions could be done to help local children. This collaboration led to the current project. The Japanese government through Japan International Cooperation Agency (JICA) has also contributed to the community by helping to establish a children's library in KV in 2015.

[9] Grades in the Philippines are referred to by G1 (1st grade), G2 (2nd grade), etc. We again follow local nomenclature.

before the afternoon school session). There was some variation in the number of days of intervention depending on the year and the type of intervention. We provide details below. Children were encouraged but not required to attend and because the interventions occurred near the children's lunch hour we also provided lunch for the treated children[10]. The drawing and storytelling interventions ran for two consecutive school years (2016-17 and 2017-18). However, the e-learning intervention ran for three consecutive school years (2016-17, 2017-18, 2018-19). This was partly because during the first intervention year (2016-17) there were issues with internet access and the e-learning intervention could not be successfully implemented so we extended it for an additional year through the 2018-19 school year.

Child educational interventions

Drawing The youngest learners (initially in G1 and G2) were provided with a drawing class. The class always began with an open-ended story and the learners were then encouraged to create their own endings to the story through their illustrations. This was designed to stimulate their creativity and cognition. This program was created by our co-author Dr. Takaaki Okumura who is an art scientist that formerly worked for Japan's Ministry of Education designing art education at primary and secondary schools in Japan. Art education in Japan is designed exactly to develop such cognitive and noncognitive skills. Staff were also given training about how to interact with the learners in order to encourage (but not distort) their creative processes. Although few studies have rigorously measured the impacts of art education on skill development, existing research finds positive effects on psychological traits and behavior (McDonald and Drey 2018) particularly for those with learning disorders (Regev and Guttmann 2005). Art therapy has also been shown to reduce anxiety in children with leukemia who were undergoing painful procedures (Favara-Scacco et al. 2001). We hypothesized this might have some parallels to a stressful high-poverty environment (Haushofer and Fehr 2014; Ridley et al. 2020). We are not aware of any other RCTs that specifically focus on the impacts of art education.

Storytelling For learners initially in G3 and G4, we worked with Mr. Rey Buhi and Ms. Jennica Gan, experts from a Ginebra Ako award-winning volunteer organization called *The Storytelling Project*, to design a storytelling intervention[11]. They provided training to staff about how to conduct the storytelling sessions in order to engage and manage learners. The goal was to improve the learners's imaginations and to instill a love of reading with the hope that this would translate into increased literacy and interest in school. The intervention involved a group reading activity in which the staff member would show the book to the learners and read out loud. Learners were again encouraged to craft their own endings to the story in order to help stimulate their imaginations. In addition to the story, the intervention also included singing and dancing at the beginning of each storytelling session to help the learners focus during the story. A meta-meta-analysis of student learning in developing countries suggests pedagogical interventions, teacher training, and improving accountability are effective (Evans and Popova 2016). Early reading interventions in developing countries can work (Graham and Kelly 2018) and our intervention shared some aspects of successful interventions such

[10] Nutrition has been shown to be an important component of school productivity (Glewwe, Jacoby, and King 2001). In that sense our intervention is a compound intervention consisting of both educational and nutrition intervention so it would be impossible to identify them separately.

[11] The Ginebra Ako awards honor individuals or organizations that have made exceptional contributions in the Philippines.

as staff training and scripted pedagogy. However, we did not find any RCTs specifically related to storytelling as way to promote reading skills.

E-learning For learners intially in G5 and above, we implemented an adaptive e-learning software for mathematics. In comparison to the drawing and storytelling that focused on broader skill development, this intervention was more focused on remediating specific mathematics deficits that might be holding the learners back. Learners were in a classroom together and each learner had their own computer to use. The e-learning software had the ability to identify and drill learner weaknesses while issues with learner focus and motivation were addressed by trained staff.

During the first year of intervention (2016-17), a Japanese edtech company called *Quipper* allowed us to use their e-learning product[12]. However, we could not implementing the program properly because of limited internet access, which is one among many difficulties that can arise in implementing RCTs in developing countries (Karlan and Appel 2017). From the middle of 2017-18 school year, we received generous support from a Japanese e-learning company called *Surala Net* including full access to their online materials and intensive training of our staff members to serve as coordinators. Internet access had also improved. Because the first-year intervention could not be implemented properly, the Surala e-learning program was extended an additional year through the 2018-19 school year. We also invited learners for two hours per week twice instead of one hour per week[13].

In general the literature on computer aided instruction finds mixed evidence of effectiveness except for a clear pattern of some impacts on mathematics scores in developing countries (McEwan 2015; Bulman and Fairlie 2016; Escueta et al. 2020). However, even this evidence produces some variation in impact depending on whether the e-learning is integrated into the curriculum, held out of school, and on the initial ability level of the students (Linden 2008; Lai et al. 2013, 2015; Mo, Zhang, Wang, et al. 2014; Mo, Zhang, Luo, et al. 2014; Mo et al. 2015; Bai et al. 2016). There have been calls for more research to help isolate mechanisms and to understand generalizability of these e-learning programs (Escueta et al. 2020).

Parent information intervention

For one of the treatment arms, in addition to the educational intervention, we also provided an information intervention for parents. Some research has found that providing information can substantially change the investment behavior of households in developing countries (Nguyen 2008; Jensen 2010). This is particularly intriguing for policymakers because information interventions can in principle be done quite cheaply, possibly at scale, and information can also spread across social networks, which can lead to *seeding* strategies (Banerjee et al. 2013). Although the impacts of information interventions tend to be small (McEwan 2015; Escueta et al. 2020), their low cost implies that even modest impacts can pass cost-benefit analyses.

Our information intervention was held for one hour in groups of 20-40 parents, which was so they would feel comfortable to ask questions. This information intervention was conducted before the start of the educational interventions. In the intervention, we stressed the importance of education to parents by providing information on how graduating from college could increase monthly earnings, on children's skill formation and the

[12] Quipper saw huge early-stage growth in the Philippines during the pandemic (Ignacio 2022).

[13] We had difficulty increasing the attendance rate particularly among learners in higher grades. Many had anxiety about mathematics and tried to avoid the e-learning program. But those who tried it often became strongly attached and continued to attend.

importance of early investments, some basic information about the costs of college and the availability of scholarships, some calculations on how *not* spending on gambling and smoking could compound over time, and finally we encouraged savings for educational investment by giving parents an *alkansya*[14]. We distributed a pamphlet biweekly for the duration of the treatment to remind the parents about the information. The education only intervention group also received an explanation of the importance of cognitive and noncognitive skills. All groups (including control) received an introduction to the project team, an explanation of the survey, assistance in filling out survey, consent, and waiver forms, and an introduction to the JICA-funded child library run by Salt Payatas.

From our pre-intervention survey, parents seemed to drastically underestimate the monetary returns to schooling. We asked parents about their expectations for their child's income at various hypothetical levels of education. We then compared the averages of these expected incomes with the actual incomes collected in our baseline survey[15]. Parents typically perceived that incomes were 20 - 60% lower than actual incomes[16].

In addition, preliminary calculations suggested that some simple financial adjustments could easily put college within reach for many of the households, which is consistent with some of the nudge and behavioral economics literature (Duflo and Banerjee 2011; Kremer, Rao, and Schilbach 2019). Although information treatments tend to have smaller effect size for self-reported attitudes and behavioral measures than for belief updating (Haaland, Roth, and Wohlfart 2023), our hypothesis was that there might be a complementarity between the parental information intervention and the educational interventions that could be used to leverage the impact. To our knowledge there are no studies looking at such interactions.

Data

Sampling, randomization, and stratification

We did our initial sample selection and survey in 2015-16 in the year before the intervention, which began during the 2016-17 academic year among students in G1-G7. We sampled at the classroom level by randomly choosing one classroom from each grade at the three schools. The randomization then occurred at the individual learner level so that some learners within the same class would be assigned to one of the two treatments or to control. This sampling was done at the classroom level to minimize disruption to teachers and administrators. We also stratified in order to sample a "star" section, which consisted of more advanced learners. In practice, there was usually only one star section per grade so this "sampling" was primarily nonstochastic.

Summary statistics

Our data come from three sources. First, children were given direct assessments, measurements, or asked survey questions. Second, we also interviewed parents or guardians

[14] Alkansya is the Tagalog word for piggy bank. This population is largely unbanked and lacks access to various financial services. We also noticed other type of savings commitment devices in other contexts in the Philippines like for tricycle drivers who sometimes have a locked alkansya where they can deposit fares for savings purposes.

[15] Because there are few college graduates in our sample of parents, we extrapolated the income for high school graduates using the returns to tertiary education reported by Montenegro and Patrinos (2014).

[16] It is somewhat unclear how such large misperceptions can persist. Financial illiteracy may play a role. In addition, respondents are largely unbanked so financial information sharing may be limited due to security concerns.

of the children. Third, we received data from teachers and principals on the children's school attendance, performance in school, and health status.

Table 1 shows mean and standard deviation for the outcomes variables in our analysis. These data are pooled across time and grouped by theme. We examined a wide variety of educational inputs and outcomes: study time, reading habits and attitude, college aspirations, self-esteem, grit, impulsivity, family environment, IQ, school outcomes, discounting, and mathematics. At baseline in 2016, we collected data from 1103 learners and this gradually decreased to 789 for the final round in 2020[17]. This population is somewhat transient with families often relocating because of labor market opportunities. However, the attrition does not look different by treatment status. In Table 1, the sample size across outcomes shows the extent of missingness and attrition by year and outcome. School outcomes were only available through 2018 and in general have the most missing data mainly because of difficulty obtaining records. Draw-a-person was not collected in 2019 due to a funding issue, and we also added Raven's Progressive Matrices beginning in 2018. A "marshmallow test", which we used to back out an estimate of the discount rate beta, was not collected at baseline.

The data indicate that the children spend more than an hour studying on both weekdays (80.88 minutes) and weekends (66.11 minutes). They also report reading 3.66 books last week and 6.39 books last month, which suggests some recall bias. 68% of children report reading after class. Among the children, 82% perceive that they are likely to go to college. This is much higher than the actual college completion rate in the Philippines, which is 16% for those 25 and above[18]. This suggests either a very low college completion rate among enrollees or perhaps some information bias in their predictions.

The next five measures used scales common in the psychology literature. Most of the measurement items are Likert scales and we re-orient the item responses to have the same direction before either averaging or summing the items depending on which is scale is used in the literature. We try to give a qualitative interpretation and also compare to measurements reported in the literature.

Reading motivation comes from the 9-item reading subscale of the *Elementary School Motivation Scale* (Guay et al. 2010) and the average of 3.03 corresponds to a "sometimes yes" interest in reading. Impulsivity is the 8-item *Domain-Specific Impulsivity Scale for Children* which aggregates a child self-report about the frequency of different types of behavior that the child engages in at school and home. Higher scores correspond to better behavior. The mean of 3.36 indicates that on average the children engage in various types of misbehavior 2-3 times a month. Family environment is the 18-item questionnaire Family Environment Scale (FES) which measures the social-environmental characteristics of the child's family (Moos and Moos 1976). Each item has a yes or no response, which we reoriented, coded as 0/1, and averaged. Values closer to 1 indicate a better family environment. The mean was 0.77. Self-esteem is the 10-item *Rosenberg self-esteem scale* that asks the children about their feelings towards themselves (Rosenberg 1965). The average of 2.92 indicates that on average the children "agree" with the items describing high self-esteem. Grit is the 8-item *Short Grit Scale* (Duckworth and Quinn 2009). Higher values correspond to more grit. Interestingly the average 3.50 in

[17] During the 2015-16 school year prior to the RCT, we originally collected data from 1441 learners in our survey. However, at the time of randomization, just prior to the 2016-17 school year, we could only locate 1103 learners. The randomization was done with this smaller population of learners. As always there is some selection into participation in the RCT. However, none of the baseline covariates seem to predict this participation decision so at least based on observables it appears to be random. Also, the child self-assessment at baseline was deliberately restricted to older children due to concerns that the questions were too challenging for younger children. This led to the smaller sample size at baseline for most of the child assessment outcomes.

[18] Authors' calculations, 2010 Philippines census (IPUMS 2020).

our sample is actually quite similar to samples from the US (Duckworth and Quinn 2009).

We used two IQ related measures. The first is the "draw-a-person" assessment which asks children to draw the figure of a person. The measure has 50 items related to the complexity of the figure drawn and it was designed to be a systematic way to measure the "nature and organization of the child's mental processes" (Goodenough 1926). Our sample averaged 29.03, which is remarkably similar to the averages reported by Goodenough (1926) for American school children in the 1920s. From 2018, we also collected a more standard IQ measure in the form of Raven's progressive matrices. We used the 12-item short form and the mean of the sum score is 6.79, which is smaller but similar to statistics reported in the literature (Arthur Jr and Day 1994). The difference is likely accounted for by the younger age of our study participants. Interestingly the correlation between our Raven's and draw-a-person measures was 0.15, which is consistent with other research questioning the validity of draw-a-person as a measure of IQ (Imuta et al. 2013).

The school outcomes (school attendance, BMI, and GPA) come from the school records and not from surveys. These data are more likely to be missing. While school attendance is quite high at 97.72%, children tend to be underweight with an average BMI of only 15.32. GPA is measured on a 100 point scale with a mean of 81.51 in our sample. Below 75 is considered failing. However, failing is somewhat rare, as only 3% of students in our data appear to be failing.

We estimated discount factors using a "marshmallow test" except with chocolate[19]. We asked children whether they preferred 5 chocolates today or 6 chocolates tomorrow, 5 today or 8 tomorrow, and 5 today or 10 tomorrow[20]. The responses were used to construct the midpoint of a discount factor that would rationalize their choices. The mean beta is 0.81. We also asked the children the same questions except in one week or in 8 days. This was intended to measure hyperbolic discounting. The mean *beta future* is 0.79 which is similar magnitude and only 11.2% of children had a beta future larger than beta, which interestingly suggests that the vast majority of these children are not hyperbolic discounters.

Finally, we implemented two mathematics assessments. The first we call CEM after the company that administered the assessment (Center for Educational Measurement). This was collected at baseline, each midline, and at endline. The second assessment (Surala) was only given in 2020 and was specifically designed to be aligned with the e-learning program from Surala Net. For both assessments we report the percentage correct, which averaged 43.16% for CEM and 79.03% for Surala.

Some of the measures we implemented after not seeing impacts in outcomes measured in the first year. Although this can be seen as searching over the outcome space, we also were not sure why we were not seeing impacts and if perhaps our measures were not well-aligned with the intervention. This, for example, occurred with the CEM mathematics assessment. At the first midline point in 2017, learners only got 36% of answers correct from which we ascertained that the assessment was too difficult for the learners and unlikely to capture any potential impacts of our interventions. After asking the company to make the test easier, the average increased to 50 in 2018. However, the test also seemed misaligned with our intervention so in the final year we also added an assessment designed specifically by Surala to be aligned with their e-learning software.

We also conducted a parental survey given that our intervention specifically targeted

[19] Our local team members suggested using chocolates instead of marshmallows because of learner preferences.

[20] Children were told they would receive chocolates according to their response to one of the randomly drawn questions. This of course raises standard issues about experimenter commitment and also the fact that the children experienced a lottery over their responses make our discount factor calculations not necessarily straightforwardly interpretable.

the parents' information set. While some of the children's items captured the household environment or the children's skills, we were also quite interested in the savings behavior of households given the encouragement for the parents to save. However, there was substantial attrition in the parental survey and lots of non-response. The respondents are largely unbanked and interviewers informed us that parents were reluctant to provide information about any savings they kept in the house for security reasons. In addition, parents usually work 6 days per week and are often away from the village so it was much easier to interview the children.

Table 1: Summary statistics

			N by Year				
	Pooled		Baseline	Midline			Endline
	MEAN	SD	2016	2017	2018	2019	2020
Learning inputs							
Minutes studying (weekday)	80.88	76.23	1072	985	774	796	752
Minutes studying (weekend)	66.11	63.69	1072	985	777	796	752
Books read (last month)	6.39	6.17	1072	986	776	795	751
Books read (last week)	3.66	4.68	1073	985	777	796	752
Read after class	0.68	0.46	1069	983	777	792	748
Child self-assessment							
College likely	0.82	0.38	1078	986	778	796	752
Reading motivation	3.03	0.47	540	986	776	796	750
Impulsivity	3.36	0.88	541	985	778	796	751
Family environment	0.77	0.13	548	983	776	795	750
Self-esteem	2.92	0.38	541	986	779	796	752
Grit	3.50	0.53	541	986	778	796	751
IQ							
Draw-a-person	29.03	9.06	1062	986	779		753
Raven's	6.79	2.92			777	796	752
School outcomes							
School attendance (%)	97.72	4.49	693	780	265		
BMI	15.32	4.59	992	472	203		
GPA	81.51	4.53	687	785	376		
Discounting							
Beta	0.81	0.20		822	706	721	643
Beta future	0.79	0.21		856	718	737	656
Mathematics							
CEM	43.16	19.83	461	400	297	312	161
Surala	79.03	16.50					340
N			1103	1015	876	838	789

Table 2 shows baseline balance for outcome variables measured at baseline as well as selected parental characteristics. We do not find any difference in the mean of the outcome variables at baseline by random assignment.

Table 2: Baseline balance

	Treatment		
	EDUCATION	EDUCATION + INFORMATION	CONTROL
Minutes studying (weekday)	73.34	69.86	70.62
Minutes studying (weekend)	70.51	75.92	72.66
Books read (last week)	4.64	4.43	4.54
Books read (last month)	6.71	6.88	7.29
Read after class	0.75	0.74	0.74
College likely	0.64	0.65	0.63
Reading motivation	3.14	3.06	3.11
Impulsivity	3.09	3.09	3.22
Family environment	0.78	0.75	0.76
Self-esteem	2.85	2.83	2.86
Grit	3.42	3.51	3.50
Draw-a-person	23.53	23.33	23.99
School attendance	97.79	97.96	97.90
BMI	16.06	15.71	15.94
GPA	81.05	80.91	81.03
CEM	32.71	31.12	33.54
Star section	0.45	0.45	0.44
Female students	0.47	0.47	0.46
Household size	3.05	2.89	3.22
Maternal education elementary school	0.17	0.15	0.15
Maternal education high school	0.56	0.63	0.60
Maternal education college	0.26	0.22	0.25
N	368	366	369

Notes: * indicates a statistically significant difference in means for the three groups at a 5% significance level.

Table 3 shows the number of intervention days and the attendance rate. The number of treatment days offered ranged between 11 and 29 and varied both by intervention and by year. The overall attendance rate among treated learners was 29% of the offered intervention days. So there were a substantial amount of "never-takers" at the intensive margin. At the extensive margin, only 22.5% of children never attended any days across all years, which shows that there is some cycling in and of out the intervention among treated children. There were no crossovers as attendance among the control group was zero. Attendance increased in the second year of the intervention and was higher for drawing and storytelling[21] E-learning was extended through 2018-19 while the other two interventions ended. All interventions then had ended by the 2019-20 school year and we collected the endline data in 2020.

[21]The days offered of e-learning in 2017-18 was lower because the Surala e-learning software was only available from November while the drawing and storytelling started from the beginning of the school year in July.

Table 3: Intervention attendance

	2016-17	2017-18	2018-19	2019-20
Days offered				
Drawing	12	25	0	0
Storytelling	12	25	0	0
E-Learning	12	11	29	0
Control	0	0	0	0
Attendance rate (%)				
Drawing	34	59	0	0
Storytelling	32	58	0	0
E-Learning	26	43	18	0
Control	0	0	0	0

Empirical strategy

We estimate the impact of our interventions using the following regression model:

$$Y_{isct} = \alpha_0^t + \alpha_1^t T_i^e + \alpha_2^t T_i^{ep} + \epsilon_{isct} \quad (1)$$

where Y_{isct} is the outcome for child i in school s in class c at time $t \in \{2017, ..., 2020\}$, T_i^e is an indicator for whether the child was assigned to the *e*ducational intervention, and T_i^{ep} is an indicator for whether the child was assigned to the combined *e*ducational and *p*arental information intervention. The parameters are superscripted by t to allow the impact of the program to differ by time period. In robustness checks, we also experimented with time invariant child/family controls X_i and baseline outcome Y_{iscb}. We also estimated the model on different subsets of the data. Testing $\alpha_1^t = 0$ can tell us about the impact of the educational intervention and testing $\alpha_1^t = \alpha_2^t$ can tell us about any differential impact of the parental information intervention. We can also examine potential dynamic impacts by examining impacts over time.

The intent to treat parameter helps us to understand the effect of the program as implemented. However, given the attendance rate, it also useful to look at a LATE parameter that identifies a weighted average along a "causal response function" (Angrist and Imbens 1995). One way to model this is to examine the impact of the attendance rate. We look at a cumulative attendance rate through period t for educational treatment (A_{it}^e) and for the combined education and parental treatment (A_i^{ep}). This variable equals 0 for control children and no-shows and ranges up to 1 for children with perfect attendance. To deal with the endogeneity of attendance, we can instrument for these two variables using random assignment T_i^c and T_i^{cp}. This gives the impact of attendance on children who are induced to change their attendance rate by assignment to the treatment. The model is given by:

$$Y_{isct} = \alpha_0 + \alpha_1 A_{it}^e + \alpha_2 A_i^{ep} + \epsilon_{isct} \quad (2)$$

In this model we again experimented with adding control and baseline outcomes as well as estimating the model on different subsets of the data.

Results

Our main results are shown in Table 4 (ITT) and Table 5 (LATE). Overall we do not find any consistent pattern of impact estimates. There are sporadic statistically significant estimates but are often the "wrong" sign, do not persist over time, and would likely disappear with adjustments for multiple comparisons[22] One exception is that there appears to be some decrease in learners' study time. This is an example of behavioral response to an exogenous intervention (Todd and Wolpin 2003; Pop-Eleches and Urquiola 2013) and is consistent with some work that finds school inputs can crowd-out private inputs (Houtenville and Conway 2008). However, the magnitude of the decrease in study time is outweighed by the 60-120 minutes of intervention, which shows that some of the transfer indeeds "sticks" to the child (Jacoby 2002).

[22]We have colored positive statistically significant impacts in blue and negative ones in orange.

Table 4: Impact estimates (ITT)

	Learning inputs					Child self-assessment						IQ		School outcomes			Discounting		Math	
	STUDYING (WEEKDAY)	STUDYING (WEEKEND)	BOOKS (LAST MONTH)	BOOKS (LAST WEEK)	READ AFTER CLASS	READING MOTIVATION	IMPULSIVITY	COLLEGE	FAMILY ENVIRONMEN˜	SELF-ESTEEM	GRIT	DRAW-A-PERSON	RAVEN'S	ATTENDANCE (%)	BMI	GPA	BETA	BETA FUTURE	CEM	SURALA
2017 endline																				
Education	-9.54 (5.65)	-1.10 (5.00)	-0.93 (0.53)	-0.082 (0.42)	-0.001 (0.036)	**-0.083*** **(0.035)**	-0.10 (0.067)	-0.039 (0.027)	-0.016 (0.010)	0.014 (0.031)	-0.078 (0.042)	**1.33*** **(0.62)**		-0.11 (0.24)	-1.50 (0.77)	-0.37 (0.42)	-0.013 (0.018)	-0.015 (0.018)	-0.96 (1.79)	
Education + Information	-8.50 (5.62)	-3.29 (4.97)	-0.47 (0.52)	0.10 (0.42)	0.009 (0.036)	-0.065 (0.035)	-0.023 (0.066)	-0.005 (0.027)	**-0.022*** **(0.010)**	-0.011 (0.030)	-0.036 (0.042)	0.31 (0.62)		-0.011 (0.24)	0.31 (0.76)	-0.030 (0.42)	-0.009 (0.018)	-0.024 (0.018)	-1.07 (1.78)	
Control mean	82.41 (73.21)	70.50 (65.09)	7.34 (7.02)	4.05 (5.38)	0.69 (0.46)	3.13 (0.42)	3.46 (0.84)	0.88 (0.33)	0.76 (0.13)	2.87 (0.39)	3.53 (0.54)	32.66 (8.67)		97.94 (2.89)	13.93 (6.60)	82.21 (4.90)	0.80 (0.21)	0.78 (0.21)	36.90 (15.73)	
2018 endline																				
Education	-8.66 (7.21)	**-11.98*** **(5.59)**	-0.52 (0.54)	-0.48 (0.39)	-0.008 (0.042)	-0.040 (0.043)	0.015 (0.079)	0.030 (0.029)	0.006 (0.011)	0.033 (0.034)	-0.067 (0.047)	-0.95 (0.73)	-0.46 (0.27)	1.05 (1.22)	-0.032 (0.54)	-0.52 (0.57)	-0.035 (0.020)	-0.035 (0.019)	-0.58 (2.90)	
Education + Information	0.52 (7.08)	**-15.72*** **(5.50)**	-0.16 (0.53)	-0.20 (0.39)	-0.019 (0.042)	-0.048 (0.042)	-0.012 (0.077)	0.041 (0.028)	0.000 (0.011)	0.015 (0.034)	-0.087 (0.046)	-0.57 (0.72)	-0.13 (0.26)	-0.36 (1.18)	0.12 (0.54)	0.014 (0.56)	0.006 (0.019)	0.014 (0.019)	-1.51 (2.87)	
Control mean	91.76 (80.24)	75.00 (65.61)	6.73 (5.82)	3.79 (4.51)	0.67 (0.47)	3.09 (0.48)	3.40 (0.91)	0.85 (0.35)	0.77 (0.12)	2.94 (0.38)	3.58 (0.52)	32.30 (7.94)	6.10 (3.09)	96.55 (7.68)	16.57 (3.34)	81.43 (4.92)	0.81 (0.20)	0.79 (0.20)	50.88 (21.69)	
2019 endline																				
Education	3.88 (7.09)	**-11.18*** **(5.05)**	-0.84 (0.49)	-0.30 (0.30)	-0.078 (0.041)	**-0.14*** **(0.041)**	-0.13 (0.077)	-0.018 (0.027)	-0.012 (0.010)	0.009 (0.033)	-0.088 (0.047)		-0.28 (0.25)				0.025 (0.018)	0.012 (0.018)	**-5.98*** **(2.76)**	
Education + Information	3.29 (7.05)	-5.74 (5.02)	-0.12 (0.49)	-0.28 (0.30)	-0.003 (0.041)	**-0.087*** **(0.041)**	-0.049 (0.076)	-0.016 (0.026)	-0.005 (0.010)	-0.008 (0.033)	**-0.11*** **(0.047)**		0.064 (0.25)				-0.004 (0.018)	-0.001 (0.018)	-3.97 (2.83)	
Control mean	87.82 (78.05)	66.35 (57.25)	6.19 (5.71)	3.24 (3.70)	0.68 (0.47)	3.06 (0.45)	3.45 (0.96)	0.91 (0.29)	0.79 (0.11)	2.96 (0.39)	3.61 (0.54)		6.97 (2.86)				0.80 (0.20)	0.79 (0.20)	53.99 (21.44)	
2020 endline																				
Education	-8.45 (6.77)	-8.57 (5.20)	-0.029 (0.45)	-0.093 (0.23)	**-0.10*** **(0.043)**	-0.022 (0.040)	-0.030 (0.078)	-0.002 (0.027)	**-0.021*** **(0.010)**	-0.010 (0.034)	**-0.11*** **(0.047)**	-0.70 (0.80)	0.11 (0.24)				-0.009 (0.018)	-0.017 (0.018)	-2.37 (3.98)	0.62 (2.20)
Education + Information	-10.09 (6.58)	-9.42 (5.06)	0.39 (0.44)	0.21 (0.23)	**-0.10*** **(0.042)**	-0.035 (0.038)	-0.026 (0.076)	-0.011 (0.026)	-0.016 (0.010)	-0.007 (0.033)	-0.090 (0.046)	-0.62 (0.78)	0.24 (0.23)				-0.005 (0.017)	-0.012 (0.017)	-1.00 (4.18)	-0.19 (2.23)
Control mean	88.15 (77.23)	64.52 (59.92)	5.25 (4.63)	2.59 (2.44)	0.71 (0.45)	2.96 (0.41)	3.41 (0.89)	0.91 (0.29)	0.81 (0.10)	2.95 (0.37)	3.53 (0.50)	28.76 (8.52)	7.46 (2.79)				0.84 (0.17)	0.83 (0.17)	64.30 (21.65)	78.87 (17.10)

Notes: For impact estimates, standard errors are reported below in parentheses. Control mean also shows the standard deviation below in parentheses. * indicates a statistically significant impact at the 5% level. Positive effects are colored blue and negative effects are colored yellow.

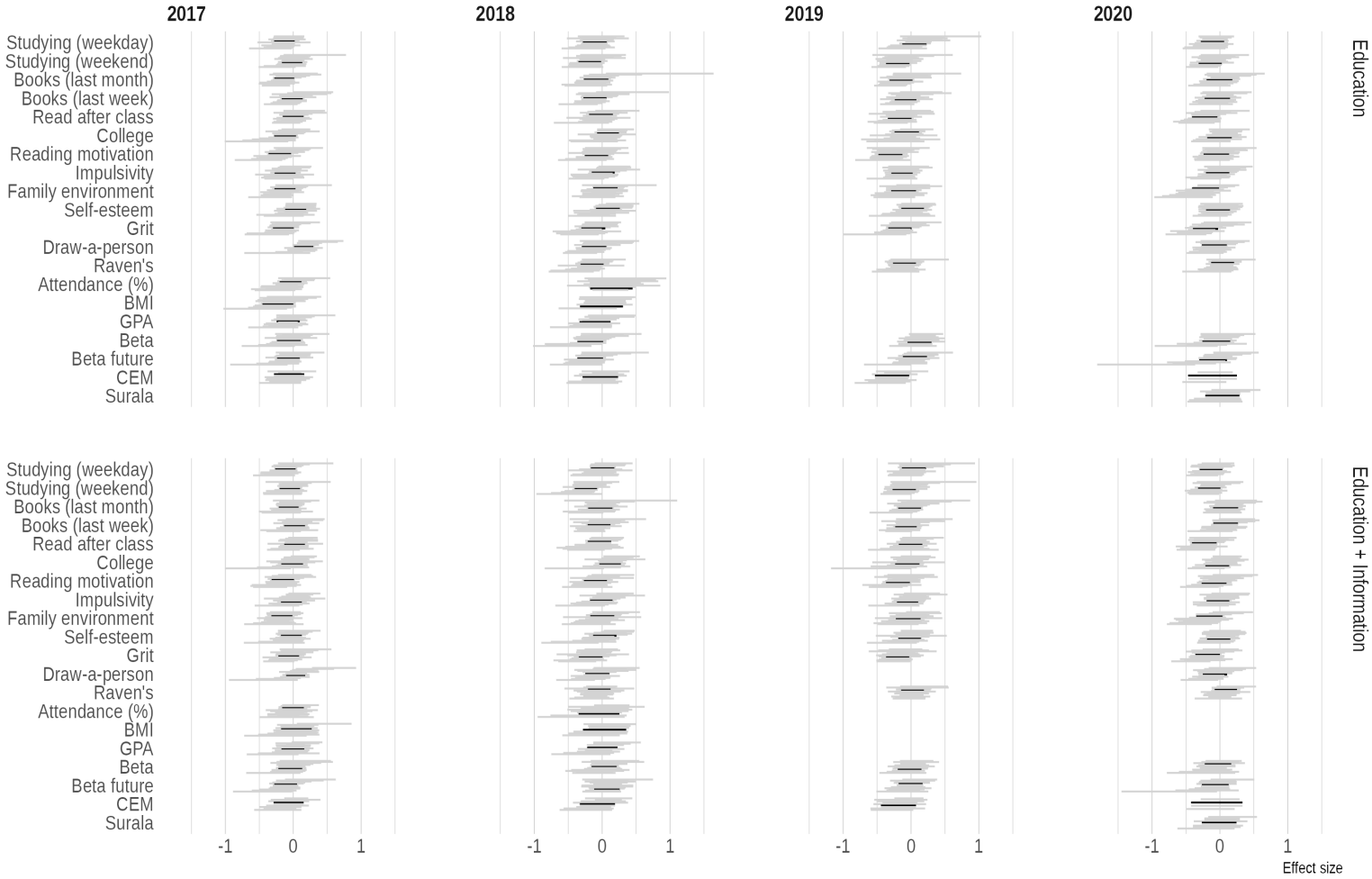
Table 5: Impact estimates (LATE)

	Learning inputs					Child self-assessment						IQ		School outcomes			Discounting		Math	
	STUDYING (WEEKDAY)	STUDYING (WEEKEND)	BOOKS (LAST MONTH)	BOOKS (LAST WEEK)	READ AFTER CLASS	READING MOTIVATION	IMPULSIVITY	COLLEGE	FAMILY ENVIRONMEN˜	SELF-ESTEEM	GRIT	DRAW-A-PERSON	RAVEN'S	ATTENDANCE (%)	BMI	GPA	BETA	BETA FUTURE	CEM	SURALA
2017 endline																				
Education	-32.55	-3.76	-3.16	-0.28	-0.003	**-0.28***	-0.35	-0.13	-0.055	0.046	-0.27	**4.55***		-0.42	-4.62	-1.41	-0.046	-0.052	-3.21	
	(19.48)	(16.88)	(1.83)	(1.46)	(0.12)	**(0.12)**	(0.23)	(0.093)	(0.034)	(0.10)	(0.14)	**(2.17)**		(0.87)	(2.46)	(1.63)	(0.063)	(0.062)	(6.30)	
Education + Information	-28.64	-11.10	-1.60	0.35	0.030	-0.22	-0.077	-0.018	**-0.073***	-0.036	-0.12	1.04		-0.039	0.90	-0.11	-0.031	-0.085	-4.02	
	(19.35)	(17.16)	(1.78)	(1.42)	(0.12)	(0.12)	(0.22)	(0.087)	**(0.036)**	(0.10)	(0.15)	(2.14)		(0.88)	(2.18)	(1.59)	(0.062)	(0.064)	(6.59)	
Control mean	82.41	70.50	7.34	4.05	0.69	3.13	3.46	0.88	0.76	2.87	3.53	32.66		97.94	13.93	82.21	0.80	0.78	36.90	
	(73.21)	(65.09)	(7.02)	(5.38)	(0.46)	(0.42)	(0.84)	(0.33)	(0.13)	(0.39)	(0.54)	(8.67)		(2.89)	(6.60)	(4.90)	(0.21)	(0.21)	(15.73)	
2018 endline																				
Education	-19.97	**-27.30***	-1.04	-1.14	-0.025	-0.10	0.079	0.071	0.013	0.075	-0.16	-2.01	-0.98	2.38	-0.27	-1.16	-0.077	-0.084	-1.64	
	(16.63)	**(13.62)**	(1.22)	(0.91)	(0.098)	(0.10)	(0.18)	(0.070)	(0.026)	(0.078)	(0.11)	(1.68)	(0.63)	(2.10)	(1.01)	(1.48)	(0.047)	(0.046)	(7.57)	
Education + Information	3.09	**-31.99***	-0.029	-0.056	-0.028	-0.12	-0.006	0.059	-0.007	0.017	-0.20	-1.73	-0.16	-0.73	0.20	0.030	0.007	0.028	-3.59	
	(16.29)	**(12.57)**	(1.23)	(0.92)	(0.094)	(0.096)	(0.18)	(0.067)	(0.026)	(0.078)	(0.10)	(1.62)	(0.60)	(2.75)	(0.94)	(1.19)	(0.041)	(0.043)	(7.67)	
Control mean	91.76	75.00	6.73	3.79	0.67	3.09	3.40	0.85	0.77	2.94	3.58	32.30	6.10	96.55	16.57	81.43	0.81	0.79	50.88	
	(80.24)	(65.61)	(5.82)	(4.51)	(0.47)	(0.48)	(0.91)	(0.35)	(0.12)	(0.38)	(0.52)	(7.94)	(3.09)	(7.68)	(3.34)	(4.92)	(0.20)	(0.20)	(21.69)	
2019 endline																				
Education	10.07	**-30.92***	-2.46	-0.71	-0.22	**-0.40***	-0.40	-0.055	-0.038	0.017	**-0.28***		-0.73				0.080	0.037	**-14.74***	
	(21.02)	**(14.97)**	(1.48)	(0.92)	(0.13)	**(0.12)**	(0.23)	(0.078)	(0.031)	(0.098)	**(0.14)**		(0.73)				(0.051)	(0.051)	**(7.33)**	
Education + Information	15.14	-18.83	-0.42	-0.60	-0.051	-0.24	-0.044	-0.059	-0.018	-0.051	**-0.36***		0.098				0.000	0.000	-11.28	
	(20.60)	(14.90)	(1.45)	(0.90)	(0.12)	(0.12)	(0.23)	(0.077)	(0.030)	(0.097)	**(0.14)**		(0.73)				(0.055)	(0.054)	(8.76)	
Control mean	87.82	66.35	6.19	3.24	0.68	3.06	3.45	0.91	0.79	2.96	3.61		6.97				0.80	0.79	53.99	
	(78.05)	(57.25)	(5.71)	(3.70)	(0.47)	(0.45)	(0.96)	(0.29)	(0.11)	(0.39)	(0.54)		(2.86)				(0.20)	(0.20)	(21.44)	
2020 endline																				
Education	-18.06	-19.92	0.14	-0.25	-0.24	-0.024	-0.062	-0.001	**-0.064***	-0.040	-0.24	-1.35	0.051				-0.009	-0.026	-5.45	2.05
	(18.37)	(14.51)	(1.24)	(0.63)	(0.12)	(0.10)	(0.23)	(0.073)	**(0.028)**	(0.090)	(0.13)	(2.17)	(0.66)				(0.048)	(0.047)	(8.40)	(6.62)
Education + Information	-26.76	-23.12	0.23	0.16	**-0.36***	-0.15	-0.087	-0.015	**-0.061***	-0.067	**-0.30***	-1.77	0.41				0.002	-0.021	-2.84	-0.65
	(19.57)	(15.68)	(1.28)	(0.66)	**(0.13)**	(0.12)	(0.22)	(0.077)	**(0.030)**	(0.10)	**(0.14)**	(2.37)	(0.70)				(0.048)	(0.048)	(11.69)	(8.11)
Control mean	88.15	64.52	5.25	2.59	0.71	2.96	3.41	0.91	0.81	2.95	3.53	28.76	7.46				0.84	0.83	64.30	78.87
	(77.23)	(59.92)	(4.63)	(2.44)	(0.45)	(0.41)	(0.89)	(0.29)	(0.10)	(0.37)	(0.50)	(8.52)	(2.79)				(0.17)	(0.17)	(21.65)	(17.10)

Notes: For impact estimates, standard errors are reported below in parentheses. Control mean also shows the standard deviation below in parentheses. * indicates a statistically significant impact at the 5% level. Positive effects are colored blue and negative effects are colored yellow. First-stage F-stats range between 36.5 and 32342.9, which varies because of missing data across outcomes and year.

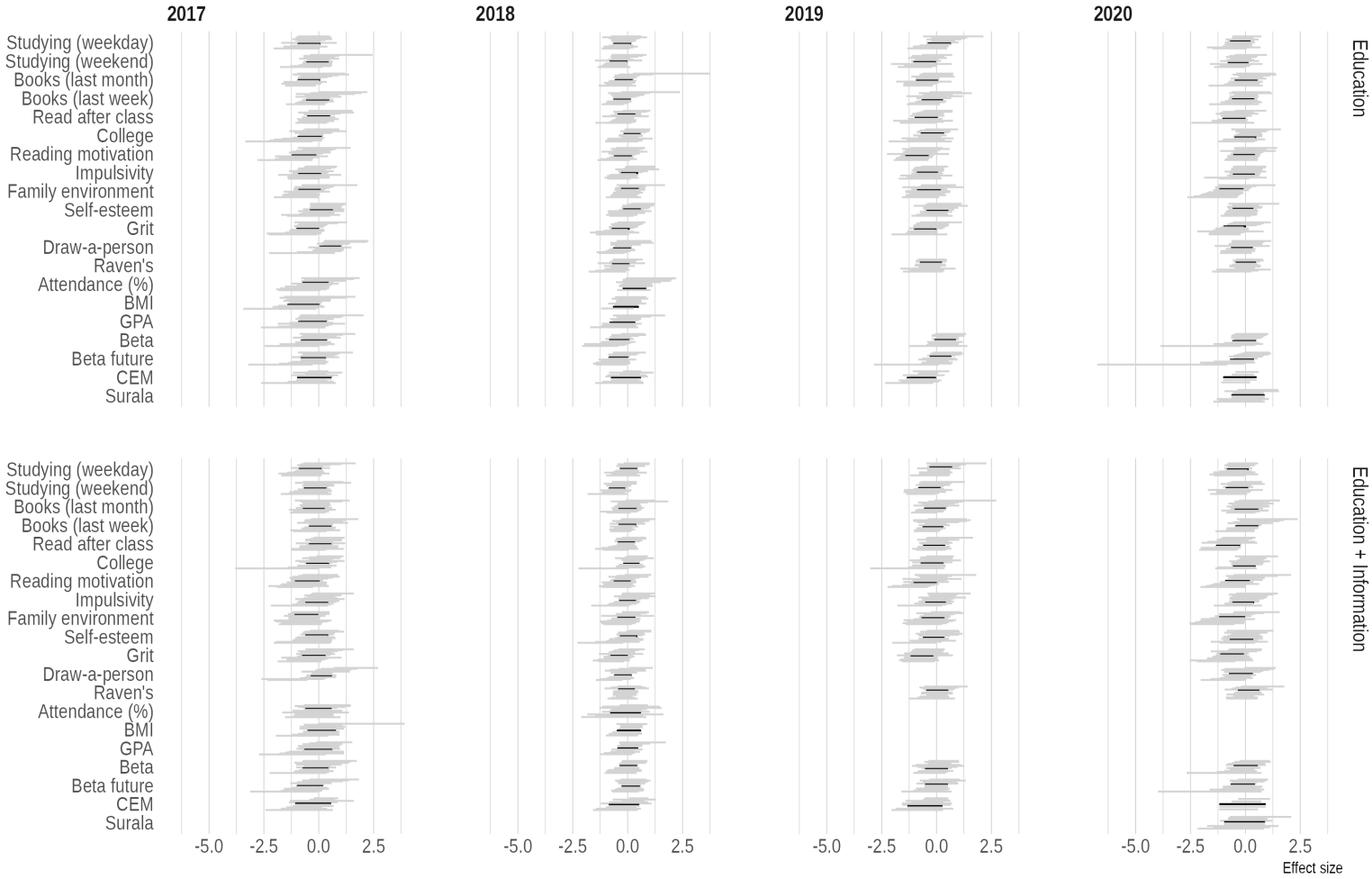
We tried several robustness checks including adding controls for baseline household characteristics, baseline learner characteristics, and baseline outcome variables. We also examined impact heterogeneity by intervention type (drawing, storytelling, or e-learning), child gender, maternal education, and star section. Figure 1 summarizes the various ITT estimates using a Forest plot that is faceted by year and treatment. The dark lines represent 95% confidence intervals centered around the impact estimates reported in Table 3. Each light gray line represents an alternative specification or subgroup analysis detailed above. As can be seen, we did not find any consistent pattern of impacts across any of the analyses. In Figure 2, we repeat the same procedure using our LATE estimates and again find no discernible pattern of impacts.

Figure 1: effect sizes (ITT) heterogeneity and robustness checks



Note: Impact estimates are colored black and robustness checks are colored light grey. The bars represent 95% confidence intervals centered around the point estin

Figure 2: effect size (LATE) heterogeneity and robustness checks



Note: Impact estimates are colored black and robustness checks are colored light grey. The bars represent 95% confidence intervals centered around the point estin

Discussion

The RCT is somewhat silent on the cause of these null results as the possibilities are multitudinous. It could have been related to the time input of the intervention or to the attendance rate. However, IV models that identify a LATE parameter for learners who were affected by the intervention also did not show any impact[23]. Perhaps it would have been more effective to incorporate the activities directly into the classrooms as extracurricular interventions were not necessarily well integrated with the school curriculum. However, curricula are more difficult to change especially in a randomized way and so offer less discretion for researchers to experiment with. Our extracurricular intervention may have also introduced some social issues for the children, especially as the randomization was done *within* classrooms (e.g., why are some children attending a special class?, treated children wouldn't get to immediately go play after school or they'd have to come to school early, etc.). Such spillovers would violate SUTVA. For example, the within-classroom randomization could have motivated control children to work harder and nullify any potential positive effects of the intervention. Alternatively, perhaps the treated children did not want to appear different than their control group classmates so that intervention itself may have been technically effective but the social context rendered it ineffective. Yet another possibility is there may have not been good alignment between the intervention and the skills of the students. The first-year mathematics assessment was too difficult for the students, which seemed to intimidate some students. Learners also had the opportunity to participate in other extracurricular programs, which may have introduced *substitution bias* into our experimental evaluation (Heckman and Smith 1995; Kline and Walters 2016). Another vexing possibility is that strong complementaries in the education production function might generate an O-ring scenario (Kremer 1993) that makes learning about the production function through experimentation difficult. This would also call for a "big push" policy (Rosenstein-Rodan 1943) that locally searching over the policy space via an RCT would be unlikely to uncover. At least in the first year, we encountered some technical issues with the e-learning technology. Such issues are likely particularly salient in developing countries and could prevent technology adoption and experimentation by schools.

Conclusion

Overall we do not find any impact of our interventions on cognitive or noncognitive skills of the learners. Particularly for e-learning and mathematics in developing countries this is in contrast to the existing literature (Bulman and Fairlie 2016; Escueta et al. 2020). We also did not find any impact of the information intervention on investment behavior, skill development, or expectations. Previous research on information interventions in developing countries is quite limited and successes might depend heavily on context. Research also suggests that nudges do not necessarily scale or generalize well (DellaVigna and Linos 2022). The drawing and storytelling interventions designed to stimulate creativity and motivation also did not show any effects even on measures highly aligned with the intervention. We adjusted some of our outcomes measures over time in an attempt to search over the outcome space for where the intervention might be showing impacts. However, we are completely transparent in reporting impacts across all outcome measures considered in the study.

The results were disappointing to the team members, school district officials, and the NGO Salt Payatas. However, a positive aspect of the study is that we have continued our

[23] Our reading of the literature is that other e-learning interventions offered similar duration of sessions and interventions. However, some offered substantially more time inputs. This would be a good issue to explore in a meta-analysis.

collaboration with the school district and with Salt Payatas as we work to identify areas where the research team can complement the efforts of local stakeholders through data analysis and capacity building. We also conducted several novel interventions related to drawing, storytelling, and examined complementaries between information and education interventions. The research identified several potential bottlenecks including resource constraints such as internet access. By reporting these results we hope to avoid publication bias on educational and information interventions in developing countries which can distort the literature (Brodeur, Cook, and Heyes 2020; Chopra et al. 2024). Rigorously conducted RCTs, regardless of findings, should be reported in the literature to give a complete picture of the state of knowledge.

References

Abadie, Alberto. 2020. "Statistical Nonsignificance in Empirical Economics." *American Economic Review: Insights* 2 (2): 193–208.

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. "Personality Psychology and Economics." In *Handbook of the Economics of Education*, 4:1–181.

Andrews, Isaiah, and Maximilian Kasy. 2019. "Identification of and Correction for Publication Bias." *American Economic Review* 109 (8): 2766–94.

Angrist, Joshua D, and Guido W Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430): 431–42.

Arthur Jr, Winfred, and David V Day. 1994. "Development of a Short Form for the Raven Advanced Progressive Matrices Test." *Educational and Psychological Measurement* 54 (2): 394–403.

Bai, Yu, Di Mo, Linxiu Zhang, Matthew Boswell, and Scott Rozelle. 2016. "The Impact of Integrating ICT with Teaching: Evidence from a Randomized Controlled Trial in Rural Schools in China." *Computers & Education* 96: 1–14.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2013. "The Diffusion of Microfinance." *Science* 341 (6144).

Banerjee, Abhijit V, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics* 122 (3): 1235–64.

Barker, Nathan, Gharad Bryan, Dean Karlan, Angela Ofori-Atta, and Christopher Udry. 2022. "Cognitive Behavioral Therapy Among Ghana's Rural Poor Is Effective Regardless of Baseline Mental Distress." *American Economic Review: Insights* 4 (4): 527–45.

Barrera-Osorio, Felipe, and Leigh L Linden. 2009. "The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia." *World Bank Policy Research Working Paper*, no. 4836.

Becker, Gary S, Scott Duke Kominers, Kevin M Murphy, and Jörg L Spenkuch. 2018. "A Theory of Intergenerational Mobility." *Journal of Political Economy* 126 (S1): S7–S25.

Boneva, Teodora, and Christopher Rauh. 2018. “Parental Beliefs About Returns to Educational Investments—the Later the Better?” *Journal of the European Economic Association* 16 (6): 1669–1711.

Brodeur, Abel, Nikolai Cook, and Anthony Heyes. 2020. “Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics.” *American Economic Review* 110 (11): 3634–60.

Bulman, George, and Robert W Fairlie. 2016. “Technology and Education: Computers, Software, and the Internet.” In *Handbook of the Economics of Education*, 5:239–80.

Chopra, Felix, Ingar Haaland, Christopher Roth, and Andreas Stegmann. 2024. “The Null Result Penalty.” *The Economic Journal* 134 (657): 193–219.

Cunha, Flavio, and James Heckman. 2007. “The Technology of Skill Formation.” *American Economic Review* 97 (2): 31–47.

Cunha, Flavio, James J Heckman, Lance Lochner, and Dimitriy V Masterov. 2006. “Interpreting the Evidence on Life Cycle Skill Formation.” *Handbook of the Economics of Education* 1: 697–812.

Cunha, Flavio, James J Heckman, and Susanne M Schennach. 2010. “Estimating the Technology of Cognitive and Noncognitive Skill Formation.” *Econometrica* 78 (3): 883–931.

DellaVigna, Stefano, and Elizabeth Linos. 2022. “RCTs to Scale: Comprehensive Evidence from Two Nudge Units.” *Econometrica* 90 (1): 81–116.

Dizon-Ross, Rebecca. 2019. “Parents’ Beliefs About Their Children’s Academic Ability: Implications for Educational Investments.” *American Economic Review* 109 (8): 2728–65.

Duckworth, Angela Lee, and Patrick D Quinn. 2009. “Development and Validation of the Short Grit Scale (Grit–S).” *Journal of Personality Assessment* 91 (2): 166–74.

Duflo, Esther, and Abhijit Banerjee. 2011. *Poor Economics*. Vol. 619. PublicAffairs New York, NY, USA.

Escueta, Maya, Andre Joshua Nickow, Philip Oreopoulos, and Vincent Quan. 2020. “Upgrading Education with Technology: Insights from Experimental Research.” *Journal of Economic Literature* 58 (4): 897–996.

Evans, David K, and Anna Popova. 2016. “What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews.” *The World Bank Research Observer* 31 (2): 242–70.

Favara-Scacco, Cinzia, Giuseppina Smirne, Gino Schilirò, and Andrea Di Cataldo. 2001. “Art Therapy as Support for Children with Leukemia During Painful Procedures.” *Medical and Pediatric Oncology: The Official Journal of SIOP—International Society of Pediatric Oncology (Societé Internationale d’Oncologie Pédiatrique* 36 (4): 474–80.

Glewwe, Paul, Hanan G Jacoby, and Elizabeth M King. 2001. “Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis.” *Journal of Public Economics* 81 (3): 345–68.

Glewwe, Paul, and Karthik Muralidharan. 2016. “Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications.” In *Handbook of the Economics of Education*, 5:653–743. Elsevier.

Glewwe, Paul, Celestine Siameh, Bixuan Sun, and Suzanne Wisniewski. 2021. “School Resources and Educational Outcomes in Developing Countries.” *The Routledge Handbook of the Economics of Education*, 218–52.

Goldstein, Thalia R, Stéphan Vincent-Lancrin, and Ellen Winner. 2013. *Art for Art’s Sake?: The Impact of Arts Education*. OECD.

Goodenough, Florence Laura. 1926. *Measurement of Intelligence by Drawings*. World Book Company.

Graham, Jimmy, and Sean Kelly. 2018. “How Effective Are Early Grade Reading Interventions?-a Review of the Evidence.” *A Review of the Evidence (January 3, 2018). World Bank Policy Research Working Paper*, no. 8292.

Guay, Frédéric, Julien Chanal, Catherine F Ratelle, Herbert W Marsh, Simon Larose, and Michel Boivin. 2010. “Intrinsic, Identified, and Controlled Types of Motivation for School Subjects in Young Elementary School Children.” *British Journal of Educational Psychology* 80 (4): 711–35.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. “Designing Information Provision Experiments.” *Journal of Economic Literature* 61 (1): 3–40.

Haushofer, Johannes, and Ernst Fehr. 2014. “On the Psychology of Poverty.” *Science* 344 (6186): 862–67.

Health Organization”. 1999. “Partners in Life Skills Education-Conclusions from a United Nations Inter-Agency Meeting. Geneva: World Health Organization, Department of Mental Health.” *Social Change and Mental Health Cluster. Retrieved January* 12: 2008.

Heckman, James J, and Jeffrey A Smith. 1995. “Assessing the Case for Social Experiments.” *Journal of Economic Perspectives* 9 (2): 85–110.

Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior.” *Journal of Labor Economics* 24 (3): 411–82.

Houtenville, Andrew J, and Karen Smith Conway. 2008. “Parental Effort, School Resources, and Student Achievement.” *Journal of Human Resources* 43 (2): 437–53.

Ignacio, Gerald. 2022. “Quipper Sees Increase in Users in Philippines.” *The Manila Times*, July.

Imuta, Kana, Damian Scarf, Henry Pharo, and Harlene Hayne. 2013. “Drawing a Close to the Use of Human Figure Drawings as a Projective Measure of Intelligence.” *PloS One* 8 (3): e58991.

IPUMS. 2020. “Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.3 [Dataset].” https://doi.org/https://doi.org/10.18128/D020.V7.3.

Jacoby, Hanan G. 2002. “Is There an Intrahousehold ‘Flypaper Effect’? Evidence from a School Feeding Programme.” *The Economic Journal* 112 (476): 196–221.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics* 125 (2): 515–48.

Jin, Yingxue, Makiko Nakamuro, Yasuyuki Sawada, Andrew S. Griffen, Tsend Ayush Sosorbaram, and Yukichi Mano. 2024. "Misfortune Seldom Comes Alone: Evidence from the Compounded Covid-19 and Typhoon Disasters in a Developing Country." *Manuscript*.

Karlan, Dean, and Jacob Appel. 2017. *Failing in the Field: What We Can Learn When Field Research Goes Wrong*. Princeton University Press.

Kline, Patrick, and Christopher R Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *The Quarterly Journal of Economics* 131 (4): 1795–1848.

Kremer, Michael. 1993. "The O-Ring Theory of Economic Development." *The Quarterly Journal of Economics* 108 (3): 551–75.

Kremer, Michael, Conner Brannen, and Rachel Glennerster. 2013. "The Challenge of Education and Learning in the Developing World." *Science* 340 (6130): 297–300.

Kremer, Michael, Gautam Rao, and Frank Schilbach. 2019. "Behavioral Development Economics." In *Handbook of Behavioral Economics: Applications and Foundations 1*, 2:345–458. Elsevier.

Lai, Fang, Renfu Luo, Linxiu Zhang, Xinzhe Huang, and Scott Rozelle. 2015. "Does Computer-Assisted Learning Improve Learning Outcomes? Evidence from a Randomized Experiment in Migrant Schools in Beijing." *Economics of Education Review* 47: 34–48.

Lai, Fang, Linxiu Zhang, Xiao Hu, Qinghe Qu, Yaojiang Shi, Yajie Qiao, Matthew Boswell, and Scott Rozelle. 2013. "Computer Assisted Learning as Extracurricular Tutor? Evidence from a Randomised Experiment in Rural Boarding Schools in Shaanxi." *Journal of Development Effectiveness* 5 (2): 208–31.

Linden, Leigh L. 2008. *Complement or Substitute?: The Effect of Technology on Student Achievement in India*. Citeseer.

Marx, Benjamin, Thomas Stoker, and Tavneet Suri. 2013. "The Economics of Slums in the Developing World." *Journal of Economic Perspectives* 27 (4): 187–210.

McDonald, Alex, and Nicholas StJ Drey. 2018. "Primary-School-Based Art Therapy: A Review of Controlled Studies." *International Journal of Art Therapy* 23 (1): 33–44.

McEwan, Patrick J. 2015. "Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments." *Review of Educational Research* 85 (3): 353–94.

Mo, Di, Weiming Huang, Yaojiang Shi, Linxiu Zhang, Matthew Boswell, and Scott Rozelle. 2015. "Computer Technology in Education: Evidence from a Pooled Study of Computer Assisted Learning Programs Among Rural Students in China." *China Economic Review* 36: 131–45.

Mo, Di, Linxiu Zhang, Renfu Luo, Qinghe Qu, Weiming Huang, Jiafu Wang, Yajie Qiao, Matthew Boswell, and Scott Rozelle. 2014. "Integrating Computer-Assisted Learning into a Regular Curriculum: Evidence from a Randomised Experiment in Rural Schools in Shaanxi." *Journal of Development Effectiveness* 6 (3): 300–323.

Mo, Di, Linxiu Zhang, Jiafu Wang, Weiming Huang, Yaojiang Shi, M Boswell, and S Rozelle. 2014. "The Persistence of Gains in Learning from Computer Assisted Learning (Cal): Evidence from a Randomized Experiment in Rural Schools in Shaanxi Province in China." *Unpublished Manuscript. Stanford, CA: Rural Education Action Program (REAP)*.

Montenegro, Claudio E, and Harry A Patrinos. 2014. "Comparable Estimates of Returns to Schooling Around the World." *World Bank Policy Research Working Paper*, no. 7020.

Moos, Rudolf H, and Bernice S Moos. 1976. "A Typology of Family Social Environments." *Family Process* 15 (4): 357–71.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J Ganimian. 2019. "Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India." *American Economic Review* 109 (4): 1426–60.

Nguyen, Trang. 2008. "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar." *Unpublished Manuscript* 6.

Pop-Eleches, Cristian, and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses." *American Economic Review* 103 (4): 1289–1324.

Regev, Dafna, and Joseph Guttmann. 2005. "The Psychological Benefits of Artwork: The Case of Children with Learning Disorders." *The Arts in Psychotherapy* 32 (4): 302–12.

Ridley, Matthew, Gautam Rao, Frank Schilbach, and Vikram Patel. 2020. "Poverty, Depression, and Anxiety: Causal Evidence and Mechanisms." *Science* 370 (6522): eaay0214.

Rosenberg, Morris. 1965. "Rosenberg Self-Esteem Scale." *Journal of Religion and Health*.

Rosenstein-Rodan, Paul N. 1943. "Problems of Industrialisation of Eastern and South-Eastern Europe." *The Economic Journal* 53 (210-211): 202–11.

Simpson Steele, Jamie. 2016. "Noncognitive Factors in an Elementary School-Wide Model of Arts Integration." *Journal for Learning Through the Arts* 12 (1): n1.

Snilstveit, Birte, Jennifer Stevenson, Daniel Phillips, Martina Vojtkova, Emma Gallagher, Tanja Schmidt, Hannah Jobse, Maisie Geelen, Maria Grazia Pastorello, and John Eyers. 2015. "Interventions for Improving Learning Outcomes and Access to Education in Low-and Middle-Income Countries: A Systematic Review." *3ie Systematic Review* 24.

Todd, Petra E, and Kenneth I Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *The Economic Journal* 113 (485): F3–F33.