UNIVERSITY of VIRGINIA
CURRY SCHOOL and BATTEN SCHOOL
EdPolicyWorks

Working Paper:

Educational Technology in Developing Countries: A Systematic Review

Rodriguez-Segura, Daniel[1]

The emergence of educational technology ("EdTech") in developing countries has been received as a promising avenue to address some of the most challenging policy questions within educational systems. In this paper, I review and synthesize all existing studies with credible causal identification frameworks of EdTech interventions in developing countries. While other studies review the evidence for EdTech interventions in developed countries, there is currently no equivalent study for developing contexts, in spite of the rising number of studies being produced. I classify studies into four thematic categories based on the type of EdTech intervention analyzed: (1) access to technology, (2) technology-enabled behavioral interventions, (3) improvements to instruction, and (4) self-led learning. I find that EdTech interventions centered around self-led learning and improvements to instruction are the most effective forms of EdTech at raising learning outcomes. Similarly, technology-enabled behavioral interventions are less promising for generating large effects but highly cost-effective given their typically low marginal costs. While expanding access to technology alone is not sufficient to improve learning, it is a necessary first step for other types of interventions. More broadly, the overall success of interventions rests on the thoughtful customization of the EdTech solution to the policy constraints at hand. Finally, EdTech interventions across all thematic areas can and should act as complements by leveraging their respective comparative advantages to address deficiencies within educational systems in developing countries.

---

[1]University of Virginia

Updated August 2020

EdPolicyWorks
University of Virginia
PO Box 400879
Charlottesville, VA 22904

EdPolicyWorks working papers are available for comment and discussion only. They have not been peer-reviewed.
Do not cite or quote without author permission. Working paper retrieved from:
http://curry.virginia.edu/sites/default/files/uploads/epw/72_Edtech_in_Developing_Countries_1.pdf

Acknowledgements: This is a working document. The active collection of articles for this working paper happened until September 2020. No error or exclusion from this review was ill-intended. The author would like to thank Beth Schueler, Isaac Mbiti, Ben Castleman, and Caroline Whitcomb for the guidance in the ideation, structuring, and writing of this paper. The author also thanks Andy de Barros, Lee Crawfurd, Thomas de Hoop, Melissa Adelman, Moussa Blimpo, Adrienne Lucas, Christopher Neilson, Elena Arias-Ortiz, Vincent Quan, and Radhika Bhula for their help facilitating papers, data, and resources. Please address all correspondence to dan.rodriguez@virginia.edu.

EdPolicyWorks Working Paper Series No. Enter 72. August 2020.
Available at http://curry.virginia.edu/edpolicyworks/wp
Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

I. Introduction

As technology evolves, the frontier of its potential applications also expands. The education sector is no exception to this, and technology has become an ever more basic input into the provision and growth of educational services over the past decades. With recent expansions of the education systems in many developing countries, and the accompanying lagging outcomes in terms of learning, retention, graduation rates, and socioeconomic equity, investments in educational technology or "EdTech" are regarded as a promising option to boost these outcomes. In particular, I define EdTech as any application of electrically-powered technologies in education that was not widely available to the public in previous decades. This includes, but is not limited to, the distribution of existing technology[1], the provision of devices with tailored software[2], the adaptation of existing and already-owned technologies[3], or the use of specialized software in communal computers[4]. Through this working definition, the current study attempts to capture the breadth and depth of the current landscape of EdTech in developing countries, in terms of actual products, but also markets, countries, and target populations.

Before adopting and adapting EdTech interventions, policymakers and educational stakeholders need to be informed about what kind of EdTech interventions have displayed the most promise for different outcomes, populations, and sets of circumstances. Given the wide-ranging and emerging nature of the EdTech field, locating and analyzing all the extant EdTech literature is not a trivial step for researchers and practitioners alike. As a response to this need, Escueta et al. (2020) offers a thorough example of a meta-review that surveys EdTech's effects on educational outcomes, focusing on developed countries. However, the most pressing challenges in the educational systems of developing countries look very differently from those of developed countries. For instance, while adult literacy rates in low-income countries is 63%[5], these rates are effectively universal in developed contexts. Similarly, net secondary school enrollment rate stands at only 34% in low-income countries, compared to 91% in high-income countries (World Bank)[6]. Furthermore, not only are the short-term goals very different across these two types of contexts, but the kind of EdTech intervention that could actually be deployed is very different due to issues

---

[1] For example, the laptops in Beuermann et al. (2015).
[2] For example, the tablets in Pitchford (2015).
[3] For example, the use of SMS texts in Berlinski et al. (2016) or T.V. programming in Borzekowski (2018).
[4] For example, the after-school program evaluated in Böhmer et al. (2014).
[5] World Bank Development Indicators: Literacy rate, adult total (% of people ages 15 and above), 2018.
[6] World Bank Development Indicators: School enrollment, secondary (% net), 2018.

related to access to technology and public infrastructure. As a response to all these factors, Escueta et al. (2020) mention that “after considering both literatures, we determined that the circumstances surrounding the ed-tech interventions that have so far been experimentally studied differed too greatly across developed and developing country education systems to allow for integrating findings from both in a way that would yield meaningful policy implications.” In short, the actual effectiveness and focus of successful EdTech interventions in developed countries may translate to very different results in developing countries, calling for an urgent need to understand the patterns within the EdTech literature focusing exclusively on developing countries.

In fact, the question of the effectiveness and appropriateness of EdTech as a tool to address the particular issues in developing countries is still an open one. While the relatively low levels of access to needed inputs such as electricity, the internet, and hardware might be challenges that hinder EdTech’s promise in developing countries, EdTech may also be particularly well-suited to address some of the most critical educational questions in these contexts. In particular, once these technological barriers are overcome, EdTech could be leveraged to address problems that would be too costly or resource-intensive to solve through other channels. For instance, EdTech could be adopted to address issues of appropriately-leveled education to deliver instruction and practice problems tailored at each student’s specific level. Such a challenge would be almost insurmountable with the current incentives and levels of educational resources, in contexts with already extremely high pupil-teacher ratios. EdTech could also be used to address issues of stakeholder accountability, such as with the implementation of cameras that monitor teacher absenteeism, and replace less-frequent but more-expensive school inspections. Furthermore, EdTech could be used to address some of the input shortages that many schools face. Simple handheld devices could be used to replace lacking inputs such as computers, textbooks, notebooks, teacher records, and teaching guides, as a single device could perform these functions by holding many documents at once. However, the effectiveness and cost-effectiveness of all these interventions has not been systematically reviewed, and hence remain an open empirical question.

On the other hand, EdTech could face important shortcomings both in terms of take-up and implementation in developing countries. One initial challenge is that the low levels of penetration of other technologies could hinder the level of familiarity with the platforms on which EdTech tools are deployed, and hence decrease of the effectiveness of an otherwise well-thought out intervention. Similarly, implementation of even well-designed programs could be especially difficult in areas with weak state capability. Either through explicit corruption leading to leakages of equipment and funds, or through poor executing capacity, weak state capacity may be a barrier towards fruitful investment in EdTech. The most cynical view is that if these governments have not been able to provide other basic inputs like textbooks and chalk to all schools, the extent to which they can deploy successful EdTech interventions is highly questionable.

To shed light on the promise and limitations of EdTech in developing countries, the current comprehensive review synthesizes the patterns and lessons found in the extant literature rigorously evaluated in developing countries. The search methods included thorough searches in scientific research repositories, working paper series from renowned research and international organizations, forward and backward tracing from key papers, and from all papers that were being subsequently added to the list. This review identifies 81 "core studies" across 36 low and middle-income countries since 2002[7], spanning 5 different methodologies, with 3 in 4 being randomized controlled trials. The core studies are organized and analyzed thematically across four different areas: (1) access to technology, (2) technology-enabled behavioral interventions, (3) improvements to instruction, and (4) self-led learning[8].

As a methodological choice, no meta-regressions are presented in this review, due to the vast diversity in the type of interventions, contexts, and outcomes of interests[9]. Given the relatively low number of studies within each category, and further variation in the types of treatment within each category, meta-coefficients may yield overly-averaged meta-parameters that could hide policy-relevant heterogeneity. Instead, the current review presents a mostly-qualitative description of the trends in the existing evidence within each of the four categories, along with summary online tables for all papers within the set of core studies. The research questions to be explored in this review are (1) across what particular thematic areas and outcomes of education has EdTech displayed the most promise in developing countries?, (2) for what EdTech interventions does the current literature suggest little evidence of their effectiveness?, (3) under what contextual circumstances do the different types of EdTech interventions work best in developing countries?, (4) what are the current gaps in knowledge about EdTech in developing countries?, and (5) how

[7] There was no restriction on search date. 2002 is simply the year of the earliest paper found.
[8] Escueta et al. also use the "Access to technology" and "Technology-enabled behavioral interventions" categories. Their "Computer-assisted learning" was replaced for a broader "Self-led learning", which also included their "Online learning" category. Finally, there were enough interventions in the "Improvements to instruction" category that did not neatly fit into the other categories, which also deserved a separate category.
[9] This methodological choice also follows Escueta et al. (2020).

do different cost structures and levels of cost-effectiveness influence the potential for scalability of an intervention? To tackle these questions, Section II begins by providing an overview of the current state of access to technology in developing countries, and the extent to which EdTech is already present in developing countries. Section III then provides a non-exhaustive overview of key constraining challenges in developing countries for which EdTech may be particularly well-suited to addressing. In Section IV, I provide the synthesis of existing evidence, organized across the four thematic areas. Section V concludes and lays out some frontiers and considerations regarding EdTech research and policy.

II. Why study EdTech in developing countries?

1. <u>The current landscape of EdTech in developing countries</u>

EdTech has started to play a role in the education of millions of children in developing countries. The Chinese market almost reached USD 2 billion in early 2019[10] and by some estimates, the Indian market is expected to reach this mark by 2021 (Sampson, et al., 2019). Globally, the EdTech market was valued at USD 17.7 billion in 2017, with expectations for a quick increase in value in following years[11]. In spite of the growing pace of the industry, this expansion does not reflect other important metrics such as a more egalitarian reach to all learners in developing countries, or the incorporation of rigorously-tested technologies. A recent analysis of the EdTech Hub database with EdTech firms from around the world (Crawfurd, 2020), shows that only 19 million out of over 450 million children in Africa are using any kind of EdTech. Furthermore, most of these users are concentrated around a few leading companies in a handful of countries, or around students watching educational programs on T.V. Over half of all EdTech firms serving developing countries, based on a widely-publicized database, are located in just three countries: South Africa, Kenya, and Nigeria (Figure A1).

Similarly, Crawfurd also points out that the potential market size matters for the extent to which EdTech innovation develops, as Figure A2 displays the positive correlation between young population and the number of EdTech firms by country. The potential market size could be driven by other factors such as language or household income. Developing an app that promotes early literacy in English or Kiswahili will have a much larger potential market than an app promoting

[10] Source: EdSurge. "Chinese Edtech sees $1.86B in Q1 2019, Bucking Plummeting Venture Trend" (May 27, 2019).
[11] Source: Frost & Sullivan. "Growth Opportunities in the Education Technology Market, Forecast to 2022" (December 15, 2017).

the same outcome but in Xhosa. Also, the presence of emerging purchasing power from low and middle-class families can play a determinant role in the decision to invest in an EdTech product. While countries with large populations like the Democratic Republic of the Congo, Ethiopia, or Bangladesh may benefit from investments in EdTech, the very low average household income, even for the standards of developing countries, might make it less appealing for private companies to invest in those contexts.

2. The state of technology in developing countries

Given the many avenues in which EdTech solutions can be implemented, and the broad nature of this review, it is impossible to establish an absolute threshold for the needs that households, schools, or educational systems must have met before adopting an EdTech product. However, most EdTech tools do require either access to connectivity features like electricity, internet, mobile coverage, a broadband connection, and/or access to hardware such as computers, cellphones, or tablets. Clearly, the extent to which these technologies are readily available in an area will heavily influence both the feasibility of implementing an EdTech intervention, and the kind of EdTech interventions available for policymakers to choose from.

Unsurprisingly, there are still large disparities across the world in terms of infrastructure that hinder the suitability of EdTech interventions in the most disadvantaged countries. For example, Figure A3 shows the level of access to two of the most basic inputs for EdTech interventions worldwide: electricity and internet. While most countries are approaching universal access to electricity, Sub-Saharan Africa still stands at 48%, lagging far behind 98% in Latin America and the Caribbean, and 92% in South Asia (World Bank[12]). Just in the three most populated countries in Sub-Saharan Africa, Nigeria, Ethiopia, and the Democratic Republic of the Congo, access to electricity stands at 57%, 50%, and 19%, respectively, leaving almost 175 million people without access to electricity in these three countries alone. The situation regarding the number of individuals currently unable to access the internet is even starker: only 1 in 4 people in Sub-Saharan Africa have access as of 2018, and in India alone there were 475 million people not using the internet in 2018 (World Bank[13]). These figures stand in sharp contrast with the degree of penetration of mobile phones in developing countries. Across the world, there are 106 mobile

[12] World Bank Development Indicators: Access to electricity (% of population), 2018.
[13] World Bank Development Indicators: Individuals using the Internet (% of population), 2018.

cellular subscriptions per 100 inhabitants (World Bank[14]), while in Sub-Saharan Africa and India there are still 82-87 subscriptions per 100 inhabitants.

Naturally, access to technologies is not only an issue of inequality between countries, but also within countries. While these intra-country inequalities can be ameliorated by higher levels of penetration, some of the most common inputs in EdTech interventions are still unlikely to reach the most deprived sectors of society in developing countries. In countries like Mexico or Peru, 94% households in the top income quintile have access to computers at home, while less than 10% of all households in bottom income quintile do (Rieble, et al., 2020). Moreover, it is often the case that these technologically-disadvantaged groups within each country are also those for which the educational outcomes lag the most. Illustrating this point, Figure A4 displays the positive relationship between district-level household access to electricity, and math achievement levels within six different countries. In this sense, the use of EdTech in developing countries also needs to be acutely aware of how its large-scale implementation may also exacerbate existing within-country inequalities, and how the intervention can be designed and adapted to reach the most disadvantaged sectors of society.

At the even more local level of schools, there are large gaps in access to technology across schools. While countries like New Zealand and South Korea have universal access to electricity and telephone facilities in all primary schools, only 45% of all primary schools in India have electricity. In countries like Cambodia, Nepal and Myanmar, less than 10% of all primary schools have access to electricity (UNESCO[15]). Access to internet at school is similarly sparse in certain developing countries: in countries like Sri Lanka, the Philippines, Kyrgyzstan, and Bangladesh, less than 10% of all schools have access to the internet (UNESCO[16]). Even the presence of computer hardware at the school-level is rare: in Niger and Zambia, there are over 500 students per computer in a school. In India, fewer than 20% of all schools have hardware for individual-use products (Sampson et al., 2019). Even among relatively high-performing developing countries such as Mauritius or Argentina, the ratio of students per computer is 1:20 (UNESCO[17]). These are critical considerations for the study and implementation of EdTech interventions in developing

[14] World Bank Development Indicators: Mobile cellular subscriptions (per 100 people), 2018.
[15] UNESCO Institute for Statistics: Proportion of schools with electricity and telephone communication facilities, 2012.
[16] UNESCO Institute for Statistics: Proportion of educational institutions with Internet access, by type, (primary and secondary) 2012.
[17] UNESCO Institute for Statistics: Learner-to-computer ratios (primary and secondary), 2012.

countries: EdTech program administrators will need to either assess and cater to the local supply of technological tools, or incorporate the provision of infrastructure and hardware.

III. What problems could EdTech address in developing countries?

Here I describe some of the challenges that educational systems in developing countries face to provide a contextual framework for this review. While this is not an exhaustive list of all policy challenges in the educational systems of developing countries, all of these have at least some potential of being ameliorated by well-designed EdTech interventions. More importantly, these shortcomings are potential culprits for the most common symptom of the need for improvement within educational systems in developing countries: the existence of the learning crisis, and are hence valuable targets to keep in mind during the design of an EdTech intervention. In particular, the learning crisis refers to the phenomenon that many children who are in school in developing countries do not learn much during the years they spend within these systems. This contrasts starkly with the gains achieved in recent decades in terms of enrollment and expected years of education per child. A large number of policy responses have now shifted their focus from trying to increase enrollment into systems without learning, to the improvement of learning levels in developing countries. In fact, in an effort to systematize the quantification of the learning crisis, the World Bank is now releasing a measure of "learning poverty", or the share of children at the end of primary who are still below the minimum reading proficiency[18]. Strikingly, 1 of every 2 children worldwide experience "learning poverty", and with a distribution heavily skewed towards low income countries. In West African countries like Chad, Niger, and Mauritania, learning poverty is virtually universal; in Sub-Saharan countries it is 87%, and even in middle income countries like Argentina, Brazil, or Colombia, learning poverty reaches about half of all children (World Bank[19]). Below I explain some of the potential drivers of these low achievement levels and areas where EdTech has great potential to improve education in developing countries.

[18] This number is also adjusted by the share of out-of-school children.
[19] World Bank Development Indicators: Learning poverty: Share of Children at the End-of-Primary age below minimum reading proficiency adjusted by Out-of-School Children (%), latest year available for each country.

1. The ramifications of increased enrollment

Recent decades have seen large increases in enrollment rates across the globe. The net primary school enrollment stood at 89% as of 2018 (World Bank[20]), up from 72% in 1970, and the number of pupils in primary school increased by 350 million during the same period (World Bank[21]). Although this is a positive trend, it poses two new challenges for policymakers within these educational systems. First, it presents the problem of the "last-mile-enrollments", where the last 11% of children still not enrolled in primary school may indeed by the hardest to enroll. Barriers like prohibitive school fees and materials, high opportunity cost of going to school coupled with low discount rates, and physical access to schools can present high price tags for policymakers when it comes to enrolling the most remote of students. For instance, in Tanzania in 2016, 19% of the population lived further than three miles away from a primary school, and 9% lived further than five miles away from a primary school[22]. In areas of high remoteness and low population density, a formal school may be hard to establish due to issues of teacher and principal recruitment, low potential numbers of students served by any one school, and difficulty to centrally monitor school performance. In these cases, policymakers and researchers alike will need to consider alternate solutions, potentially even drawing from EdTech if the current infrastructure allows it, to complement the currently available menu of options to increase school enrollment.

The second issue that arises from the increased enrollment rates is the pressure on the already strained school resources and personnel. For instance, Figure A5 shows that while the number of in-school children in low-income countries rapidly increased after 2000, the pupil-teacher ratio remained largely the same, displaying the system's capacity to barely catch up in terms of teacher recruitment. Similarly, Figure A6 shows just how thinly-stretched teacher capacity in many developing countries really is, as the pupil-teacher ratio in low-income countries is almost three times larger than that of high-income countries. As Duflo et al. (2011) point out, in practice, the fact that on average teachers have to deal with 40 pupils at the same time translates into a lack of bandwidth to cater to all students in their classes, and the wide distribution of achievement levels that comes with these students. This situation is worsened by external political incentives to focus on high-performers, resulting in work such as Glewwe et al. (2009) identifying

[20] World Bank Development Indicators: School enrollment, primary (% net), 2018.
[21] World Bank Development Indicators: Primary education, pupils.
[22] Author's own calculations from a still unpublished manuscript with Brian H. Kim, "Far from home: mapping education deserts in developing countries".

that teachers tend to teach to the top students within classes. Ultimately, these pressures accentuate the increased within-country, within-school, and within-class inequalities that emerge from the high number of first-generation students recently entering the system (Muralidharan et al., 2019).

2. Weak teacher knowledge and rampant absenteeism

Traditionally, teachers have been an essential input into the education production function. Their aptitude, teaching capacity, effort, and content knowledge plays an important role in student achievement (Chetty et al., 2014). Unfortunately, most of these characteristics are generally lacking among teachers within the educational systems of developing countries (Global Partnership for Education, 2019). One of the "extensive margins" around this issue is how teachers broadly spend the time during which they are supposed to be teaching. Figure A7 shows how instructional time is spent by teachers across four East African countries, and how high the prevalence of teacher absenteeism really is. This behavior is prevalent in other regions of the world too: in West African nations like Niger, Togo, and Nigeria, the teacher absenteeism rate was in the 14-18% range between 2012-2014 (SDI), and in India this number was around 24% in 2010 (Muralidharan, et al., 2010). Even when teachers are present at school, they are not always actively teaching. Even in Kenya, the best performer among the four countries shown in Figure A7 only 43% of the time teachers are expected to be teaching is spent actively engaging in class. On the other end of the spectrum, of 4 hours and 17 minutes in a full school day in Mozambique, students only get about 1 hour and 33 minutes of instruction every day. Beyond the implications for learning, this implies serious fiscal burdens on countries that are already lacking public funds. Muralidharan, et al. (2010) estimate that teacher absenteeism alone is responsible for the loss of about USD 1.5 billion per year in India. Since government expenditure in 2010 in India was about USD 102[23] per primary student, this leakage could double investments for almost 15 million students.

Another worrying pattern among teachers in developing countries follows the "intensive margin" through the lack of mastery in the content knowledge that they are expected to teach. Even when teachers are actively engaging in class, students' learning process can be hindered if the teachers themselves have gaps in their own understanding of the subject. The Service Delivery

[23] The World Bank's World Development Indicators suggest the "Government expenditure per student, primary (% of GDP per capita)" was 7.49% for India in 2010 (the year for Muralidharan et al. (2010)'s estimate), and the same source reports that GDP per capita (current US$) for India in 2010 was USD 1,358.

Indicators (SDI) data collection efforts also administered a basic test of knowledge, which comprised material from lower and upper primary school. SDI defines “mastery of minimum knowledge” as answering all questions pertaining to the grades that the teacher is in charge of (i.e. lower primary or upper primary) correctly. More leniently, Figure A8 shows the share of teachers attaining 70% of minimum proficiency. Only 2 in 3 teachers in Kenya, the best performer, achieve minimum proficiency. In the most critical case, Madagascar, less than 2% of all teachers achieve this threshold. Therefore, even if teachers are engaged in teaching, these numbers question the extent to which teachers, themselves the product of these educational systems, also possess the foundational numeracy and literacy skills they are expected to nurture in their students. In this sense, EdTech could step in as a complement or as a substitute for classroom instruction to fill in content gaps teachers may have.

In a vacuum, a potential avenue to incentivizing higher effort from teachers, and the attraction of a more talented workforce into the profession, is an increase in salaries. Yet, recent evidence suggests that this policy may not be effective for several reasons. Firstly, Evans et al. (2020) draw evidence from 15 African countries and find teachers’ hourly wages are higher than those for workers with comparable education and experience. Annual wages ranged from 1.5 times GDP per capita in the Democratic Republic of the Congo to 5.1 in Zambia. In terms of total monthly wages, Evans et al. (2020) find that in 5 of their countries, teachers are paid more than other comparable professions, less than other comparable professions in 7 other countries, and found statistically insignificant differences in 3 countries. This is suggestive evidence that teachers are not systematically underpaid in many developing countries.

Even if teacher pay is raised, this is unlikely to yield better performance if it is not accompanied by behavior-modifying incentives. For instance, a study which experimentally doubled teachers’ salary in Indonesia on a permanent basis (de Ree, et al., 2018) led to precise zero improvements in student learning. While it did increase teacher satisfaction, it is not clear that this is a binding constraint in developing contexts. For instance, between 2014-2016, 3 in 4 teachers from a nationally representative sample of schools, from grades 1-7 in Tanzania reported being satisfied or very satisfied with their current job, and with the support they got from the school, and over half reported being satisfied with their salary and the level of government support[24] (Mbiti et al, 2019a; Mbiti et al, 2019b; Mbiti et al, 2019c). As long as teacher capacity

[24] This refers to the teachers (n=998) in control school from the papers referenced above for which data on employment satisfaction exists.

and incentives are not aligned with learning, policies tackling simple input provision such as higher salaries are also unlikely to succeed. As such, EdTech can be leveraged to offset these deficiencies by either supporting teachers through innovations like lesson scripts on handheld devices, or to improve the quality of instruction by directly reaching students in the areas where teachers have content gaps.

3. Overambitious and fast-moving curricula

Beyond the tangible inputs, behaviors and incentives of education stakeholders in developing countries, another issue that has received considerable attention in recent years is the actual curriculum mandated to be taught in school. Curricula in developing countries still retain many features from those designed by colonial powers, such as the language of instruction, the pace of learning, and the subjects covered (Mwiria, 1991; Malissa and Missedja, 2019). Glewwe et al. (2009) and Mwiria (1991) highlight the explicit decision of colonial powers to create a curriculum nested within an educational framework that excluded most students through the neglect of foundational numeracy and literacy skills in favor of more vocational skills, and the inclusion of high-stakes examinations that determined students' promotion to higher levels of education. These curricular features often interplay with socioeconomically disadvantaged populations, resulting in what Pritchett and Beatty (2015) call "overambitious curricula". Overambitious curricula move faster, aim higher, and span wider, than the realistic amount of material that could be taught within the contextual constrains. These curricula tend to be scattered across several subjects, taught in colonial languages, and do not have provisions for children that fall behind. The issue of overambitious curricula is typically worsened by large class sizes that do not allow teachers to help the students who fall behind.

Significant policy and research efforts on several fronts have recently tried to address the issue of overambitious curricula. The first is the targeting of appropriately-leveled material to children, through initiatives like "Teach at the Right Level" (Banerjee et al. 2016), which started in India through Pratham and has now spread to other countries, particularly in Sub-Saharan Africa. Even when this level of customization is not possible, policymakers can narrow the mandated curriculum, particularly for the critical earlier grades while children learn foundational numeracy and literacy skills, in an effort to devote more instruction to the key skills for future academic success. For example, preliminary work by Mbiti and Rodriguez-Segura (2020) in

Tanzania studies a curricular reform which shifted a significant share of the instructional time in grades 1 and 2 from tangential topics like "vocational skills", towards basic arithmetic, reading, and writing skills. This policy shift led to improved numeracy and literacy skills. Finally, if the official curriculum could not be adjusted, policymakers could leverage EdTech's potential to customize content delivery and practice exams to more appropriately meet each student's needs.

4. Inexistent or insufficient school inputs

Perhaps the most evident issue at first sight in schools in developing countries is the lack of appropriate physical inputs like desks, books, blackboards, computers, or notebooks. Figure A9 gives a sense of the stark physical environments in schools in developing countries. In Niger and Nigeria, less than half of all students had paper to write on, and in Togo there were approximately 66 students per math textbook. Data from Mbiti et al, 2019a; Mbiti et al, 2019b; Mbiti et al, 2019c shows that virtually all schools in Tanzania had a blackboard, but only 80% had chalk, and Figure A9 shows that 3 in 10 classrooms do not allow for all students to read the blackboard properly. Even in Kenya, the best performer in terms of textbooks, there were 2.6 students per math textbook, which still complicates the logistics of sharing textbooks, and bringing them home. At a broader school-level, the actual school facilities where students congregate are similarly poorly equipped and maintained. For instance, between 2013-2016 only 1 in 5 schools in Tanzania had a library, only 1 in 2 had a water source within the school premise, and up to 1 in 3 schools had trash inside the classroom.

It is worth noting that several comprehensive reviews of evidence, seminal papers, and meta-analyses have found that interventions that simply address these input constraints through "supply-side" provisions (Masino and Niño-Zarazua, 2016; McEwan, 2015; Murnane and Ganimian, 2016; Glewwe and Muralidharan, 2016), by lowering implicit and explicit costs of schooling (such as the provision of school uniforms, as in Evans and Ngatie, 2020), or by providing better school supplies (as in Glewwe et al., 2009) do not lead to improved learning. Other work such as Sabarwal et al. (2014) shows that in contexts like Sierra Leone where the government has not consistently equipped schools with inputs like textbooks, schools will be less inclined to use these textbooks, as they foresee a volatile supply of inputs in the future. Instead, they preferred to "smooth their consumption" of textbooks by storing any books received. Therefore, these inputs should be understood as necessary but not sufficient inputs to the learning production function

(Sampson et al., 2019). For example, Mbiti et al. (2019a) show that while the provision of school grants does not lead to improved learning, when these grants are coupled with appropriate teacher incentives, the joint treatment has a much larger effect than either branch alone. In other words, while inputs themselves may not be enough to raise learning standards, they can act as augmenting complements to any learning-oriented intervention, including EdTech.

IV. Methodology for this review

1. Inclusion and exclusion criteria

The primary aim of this project is to understand the broad patterns in the existing evidence of the effectiveness of technology as a policy tool towards improved educational outcomes in developing settings. As part of the review, 81 different studies were identified, which I refer to as the "core studies". The scope of this review, and hence the inclusion of the 81 papers into the group of core studies studying EdTech in developing countries, was determined through four main inclusion parameters. These parameters were a. the quality of the evidence, b. the stage in the publication pipeline, c. the context where the study was conducted, and d. the inclusion of at least one treatment branch with a technological component that meets the definition of EdTech provided in the introduction[25].

a. Quality of evidence

While a vast number of policy reports, and rigorous descriptive and theoretical studies explore EdTech (see Rubagiza et al.,2011; Henessy et al., 2010; Trucano 2015, 2016a, 2016b; Chinn and Fairlie, 2010; Bulman and Fairlie, 2016), this review focuses on studies with causal identification strategies, which were evaluations of the effectiveness of a technological feature in education. In practice, this meant focusing on papers which reported using experimental methods (RCTs), or quasi-experimental methods, including propensity-score matching. Within these parameters, all papers were included, regardless of the quality of the actual methodological implementation or publication outlet quality. Table 1 below outlines the prevalence of studies in the set of 81 core papers, where the most salient feature is that 3 in 4 studies are randomized controlled trials.

[25] For additional details on the methodological approach, see Appendix B.

Table 1: number of studies in this review by main methodological tool

Empirical methodology	Number of studies
RCT	61
Difference-in-differences (DiD)/Trend-break	14
Propensity-score matching (PSM)	4
RD	1
IV	1
Total	81

b. Publication stage

The second feature used to filter studies is the publication stage. Following Escueta et al. (2020), the scope of the current search is to be as inclusive as possible of all existing evidence, regardless of its place along the publication pipeline. This is also the strategy that Escueta et al. (2020) follows to deal with "file drawer bias", or the tendency for studies with positive or negative results to be published in peer-reviewed journal more often than studies with null results[26]. In sum, this approach means that the current search includes published papers (47%) and working papers (43%), but also policy reports which do not currently have a full paper and unpublished manuscripts which were referenced by other documents but which were not publicly available. For the unpublished manuscripts and policy reports (10% of all core studies), the authors of the studies were contacted, and in all cases the authors gracefully agreed to share the studies to be included in the current review[27]. Finally, the breakdown of the core studies by their stage in the publication pipeline is shown below.

Table 2: number of studies in this review by stage in the publication pipeline

Publication stage	Number of studies
Published paper	38
Working paper	35
Policy report	6
Unpublished manuscript	2
Total	81

c. Context

Developing countries are at the core of this review. Therefore, the third criteria for inclusion was that all studies needed to be set in developing countries. While the definition of what constitutes a developing country may vary, this review was as inclusive of studies as possible in

[26] For a striking example of this, see DellaVigna and Linos (2020). This paper shows how published papers on applications of "nudge theory" are on average more effective than all interventions implemented by "nudge units", whether they are published or not.
[27] This review does not include projects which have not concluded, or for which not even preliminary results are publicly available, as promising as these may be. For a non-exhaustive glance at projects involving EdTech in developing countries which were not included in this review because of this reason, please see Table A1 in the appendix.

terms of context. Therefore, while countries like Zambia or Colombia are indisputably included under any definition of "developing country", countries like Russia, Chile, and Israel, may not always be. The presence of these countries follows a more traditional and inclusive classification of countries for the sake of widening the reach of this review. Similarly, the inclusion of these countries also follows the classification made by other studies. For instance, the inclusion of Angrist and Lavy (2002) and Malamud and Pop-Eleches (2011) in Israel and Romania respectively, follows the choice made in Muralidharan et al. (2019), while the inclusion of the Bettinger et al. (2020) study in Russia follows the authors' own classification of Russia as a developing country. The breakdown of studies by country is displayed in Table 3:

Table 3: number of studies in this review by country where the study took place

Country	**Number of studies**
China, India, Peru	9
Kenya	4
Chile, Colombia, Indonesia, Malawi, Nigeria, South Africa, Zambia	3
Costa Rica, Ghana, Tanzania, Uruguay	2
Angola, Botswana, Cambodia, Dominican Republic, Ecuador, El Salvador, Gambia, Haiti, Honduras, Israel, Jordan, Mexico, Niger, Pakistan, Paraguay, Romania, Russia, Rwanda, Senegal, Sudan, Uganda	1
Total	81

d. Technology

The final filtering criterion was that all studies needed to have a major component evaluating an application of technology with the goal of improving educational outcomes. Within these parameters, the search was cast as widely as possible. In other words, the focus of the study could have been a major technological intervention, such as the livestreamed instruction in Johnston and Ksoll (2017), but it could have also been a complement to a treatment branch such as the cameras and incentivized payments for teachers in Gaduh et al. (2020), or it could have also been just another experimental arm besides other non-tech arms such as the interactive boards and computer labs in Berlinski and Busso (2017). To understand the extent of the technological component within each study, Online Tables 1-4 contain information on the specific intervention in each paper, and the kind of technology used.

2. <u>Classification of studies</u>

The studies identified reflect significant diversity in the types of treatments, contexts, targeted stakeholders, and scale of interventions. In practice, this diversity had several implications for how this study was conducted. First, no attempts to conduct a formal meta-analysis with unified meta-point estimates was attempted. The main reason for this methodological choice was that any aggregate estimate of whether "EdTech works in developing countries" would just mask the crucial heterogeneity that stems from the broad definition of EdTech used for this review, the dependence of effectiveness on the context, and the targeted outcomes. Instead, the core studies were coded[28] into four broad thematic categories: "Access to technology", "Technology-enabled behavioral interventions", "Improvements to instruction", and "Self-led learning". Interestingly, Escueta et al. (2020)'s four thematic categories do not fully overlap with the categories for this review, as the type of intervention and issues addressed in the current body of literature varies greatly between developed and developing countries. In reality, studies may not neatly fit into one category or the other. For instance, an argument could be made that all the "One-Laptop-per-Child" (OLPC) interventions like Beuermann et al. (2015), Cristia et al. (2017), de Melo (2014), and Cordero-Meza (2017) were ultimately about "self-led learning" at home, not necessarily access to technology. However, given that the most proximate goal of the project was to increase children's access to technology, these were categorized as "access." The table below displays the breakdown of all core studies into the category which they were assigned for the current review, as it is also shown for each study on Online Tables 1-4.

Table 4: number of studies by the area of classification within this review

Publication stage	**Number of studies**
Access to technology	22
Technology-enabled behavioral interventions	12
Improvements to instruction	20
Self-led learning	27
Total	81

V. Findings

a. Access to technology

Much policy and research attention has been devoted to the issue of access to technology. Approximately one fourth of studies focused on access to technology. Large global inequalities in access have motivated initiatives such as the highly popular "One-laptop-per-child" (OLPC), where governments, donors and NGOs aim to have a computer-pupil ratio of one to one, either through direct provision of laptops to students or through classroom sets large enough for each child to have a laptop to themselves. Investments to increase students' access to technology at

[28] For the full coding and more detailed information on all the core studies included in the review, please see this online document.

school have also become a clear policy priority for even the lowest-income countries (Kozma and Surya Vota, 2014). In spite of the momentum to improve access to technology, the evidence is at best mixed, and realistically does not suggest that the mere provision of technological tools translates directly into higher academic achievement.

None of the evaluations of the OLPC initiatives across Latin America found significant results on scholastic outcomes (Barrera-Osorio and Linden (2009) in Colombia; Beuermann et al. (2015), Cristia et al. (2010, 2017), in Peru; de Melo et al. (2014) in Uruguay, Meza-Cordero (2017) in Costa Rica). Similarly, a long-term follow up of the OLPC in Uruguay also finds null results on educational attainment (Yanguas, 2020). Bando et al. (2017) finds that replacing regular textbooks for laptops in Honduras had no statistically-significant effect on learning, and costs about USD 48 more per student than the status quo. The only exception within the evaluation of OLPC policies is Mo et al. (2013) in China, where the authors do not find any effects on language, but find effects of 0.17 SD in math achievement, as well as an increase in the amount of time spent using an educational software. A qualitative study in Brazil (Lavinas and Veiga, 2013), not included in the set of core studies, also reviews the results of OLPC initiative in Brazil, and finds that the persistent under-utilization of the computers and lack of teacher training on how to incorporate the equipment into daily instruction hindered the potential of the project. Similarly, Barrera-Osorio and Linden (2009) find that the most problematic step is the actual incorporation of computers into the instructional process.

The presence of null results for most OLPC interventions does not necessarily imply that if students are provided with computers, they did not use them: in spite of the lack of positive effects on grades, Meza-Cordero (2017) finds that treated students with OLPC did experience an increase in the amount of time they spent using a computer, at the expense of time doing other activities like homework and outdoor activities. Indeed, studies such as Angrist and Lavy (2002)[29], and Malamud and Pop-Eleches (2011) find *negative effects* on academic outcomes as a result to the provision of technology to students. In spite of the negative to null effects on academic learning as a result of increasing access to technology, there is evidence to believe that this kind of intervention can improve computer skills and familiarity with technology. In particular, Mo et al. (2013), Bet et al. (2014), Malamud et al. (2019), Malamud and Pop-Eleches (2011), and

[29] Note that Figure 1 does not show Angrist and Lavy (2002) with a negative outcome, as Figures 1-4 plot the largest gain in any subject measured. In the case of Angrist and Lavy (2002), they find negative effects on math, and no effects on language, so the largest gain is the null result on language.

Beuermann et al. (2015) find that the exposure to technology led to an improvement in familiarity with technology, up to an increase of 0.30 SD in "digital skills" in the case of Bet et al. (2014). If digital skills are also considered a valuable outcome from this type of intervention, then there is more evidence to suggest that exposure to tools like computers naturally increases students' familiarity with technology and digital skills than there is to suggest that these technologies can raise test scores by themselves.

There were four interventions providing handheld devices, with more mixed results than the provision of computers. While Pitchford (2015), and Mensch and Haberland (2018) find positive effects of the handheld devices, Habyarimana and Sabarwal (2018) find null effects. Piper et al. (2016) find that the treatment arms providing a literacy program plus handheld devices for teachers or students were at most as effective, and less cost-effective than the base literacy program. Among these four interventions, the two with the strongest case for the use of technology, Pitchford (2015), and Mensch and Haberland (2018), also had an important element of in-person support. In the case of Pitchford (2015), teachers and volunteers supported the use of the tablets with mathematical content, and in the case of Mensch and Haberland (2018), the provision of e-readers was complemented with routine group meetings. On the other hand, a treatment branch of Habyarimana and Sabarwal (2018) included content tailored to the national curriculum, but there was no in-person support for the users of the handheld devices. These results highlight again that the mere provision of hardware may not be enough, if this is not accompanied by proper in-person pedagogy or encouragement, even if this is not one-to-one with the learner.

The most salient exceptions in terms of raising student achievement levels within the category of access to education were the three papers looking at the effect of large-scale interventions providing high-level access to technology. Specifically, these three papers were Kho et al. (2018), with the large-scale provision of internet access in public schools in Peru; Navarro-Sola (2019) in in the case of *telesecundarias*[30] in Mexico; and Seo (2017) with the electrification and provision of instruction-enhancing tools in Tanzania. All of these interventions were targeted at a much larger scale than specific individuals or schools, and consisted of helping deprived regions catch up technologically with other areas within the country, as opposed to the provision of more advanced technologies (e.g. laptops in OLPC) which are not as widespread within each

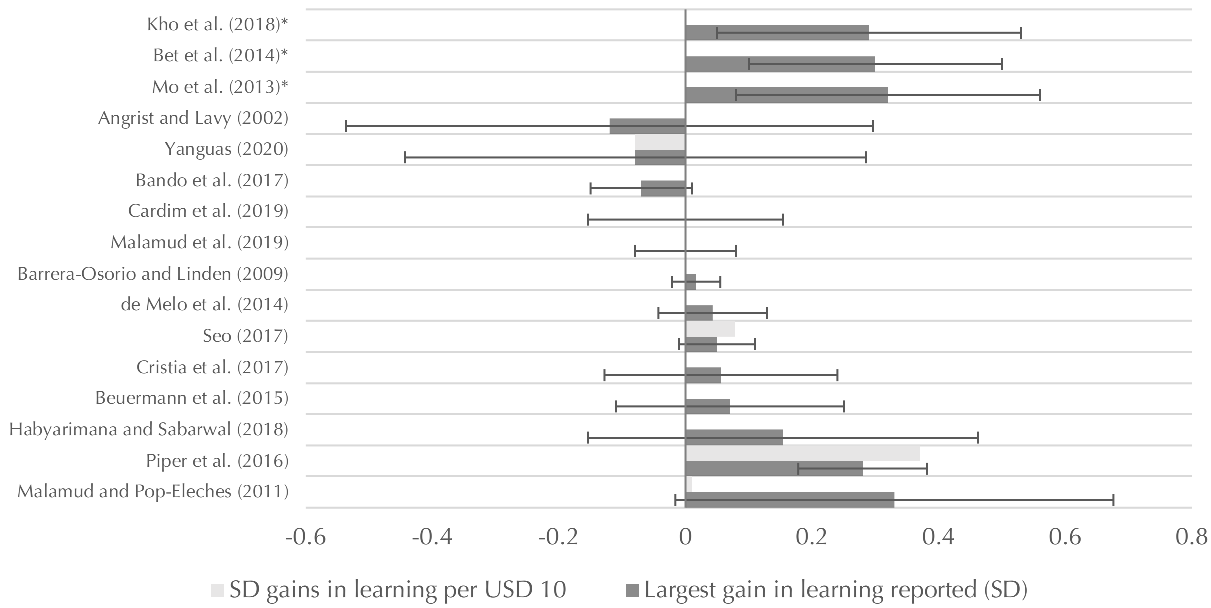
[30] According to the author of the study, "Telesecundarias are a type of junior secondary school that delivers all lessons through television broadcasts in a classroom setting, with a single support teacher per grade. The televised content follows the national curriculum and is complemented with learning guides and in-classroom work and discussions."

country. These interventions may be suggestive evidence that large-scale infrastructure-enhancing interventions in underprivileged areas may be effective in complementing students' education and narrowing within-country inequalities, such as in the case of Seo (2017) and Kho et al. (2018).

In all, it seems unlikely that the mere provision of hardware will yield to improved learning outcomes, as Sampson et al. (2019) also point out. In fact, the median effect of all the studies included in this category is an imprecise null effect. Even more importantly, when these interventions are provided at the student-level like in Angrist and Lavy (2002), or Malamud and Pop-Eleches (2011), as opposed to the mass construction of infrastructure, they also tend to also be very costly. The very low gains in learning coupled with the high price tag of these interventions should make policymakers weary of programs that simply increase access to technology, with the important exception of programs that are explicitly intended to increase digital skills. However, access to technology is a necessary, but not sufficient, requirement for the implementation of other kinds of EdTech interventions. Therefore, as long as interventions that increase access to technology are either well-accompanied by pedagogical tools, or designed as a stepping-stone for other type of interventions, they should remain in the menu of options for policymakers in some form. Finally, policymakers should still consider the trade-off of implementing interventions that increase access to technology to then implement another type of intervention, and simply designing the second intervention around more prevalent technologies such as SMS messages, phone calls to feature phones, or radio instruction[31].

[31] For instance, see Trucano (2010) for a high-level overview of radio instruction programs, and Ho and Thukral (2009) for an overview of the evidence on the effectiveness of radio instruction.

Figure 1: comparison between gains in learning reported and cost-benefit ratio for interventions in "Access to technology" category



Notes: in order to display the full potential of each intervention, "learning gain" coded as the largest gain in any field of learning, whether it is an academic subject like "math" or a less established area like "computer skills". The "SD gains in learning per USD 10" corresponds to the largest effect in any field of learning, divided by the per-pupil spending in USD, divided by 10. Studies for which authors did not report enough information to standardize gains into SD units are not in this plot. Studies denoted with a star (*) did not report enough cost information to obtain a per-pupil estimate, and hence a cost-benefit ratio. Studies without confidence intervals did not report standard errors in the results. Studies are sorted by whether they do not have cost information first, and then by the "Largest gain in learning reported" bar.

b. Technology-enabled behavioral interventions

Shaping behavior seems like a less straightforward endeavor than the provision of inputs. This requires deep knowledge about the specific constraints to be relieved, the availability of a channel through which behavior-shaping incentives can flow, and a well-designed intervention informed by a credible theory of change. Still, interventions that curb behavior are promising avenues to shape systemic issues in a cost-effective manner. In this section, I begin by reviewing interventions aimed at affecting teacher behavior, and then interventions that curb parental and student behavior.

The ingrained issue of teacher absenteeism and accountability was tackled by both Gaduh et al. (2020) in Indonesia, and Duflo et al. (2012) in Kenya by providing cameras with timestamps, and teachers were required to take frequent pictures with their students to prove that they were in school. Furthermore, both interventions conditioned at least a portion of the teacher's pay to their presence in school, as verified by the cameras. Both interventions proved effective, raising students' test scores by 0.17-0.20 SD. In the case of Gaduh et al. (2020), the treatment arm with the camera was one of the treatment arms (among others) which also sought to increase school-level accountability such as the public dissemination of scorecards. Although the camera treatment arm was the most effective at raising student outcomes, the other two treatment arms were also effective. Furthermore, there was suggestive evidence that the camera indeed led to changes in teacher behavior, emerging as a potential mechanism for the increased test scores. In spite of these successes, implementation and take-up do play a major role in the success of this kind of intervention. For instance, Adelman et al. (2015) implemented an intervention which had as one of its components a platform where teachers could send daily photographs to verify their presence, similar to Duflo et al. (2012) and Gaduh et al. (2020). The authors highlight the very low take-up of the program, and serious logistical challenges at the time of implementation, ended up hampering the effectiveness of the intervention. For instance, the authors mention that “The program faced challenges from the start, including delays and technical problems that made it hard to implement it as planned” and “There were so many problems getting schools ready for the pilot that the program ended up starting months late […] This short implementation period reduced the chance of seeing any change in teacher behavior or student learning". Therefore, even if the behavioral intervention is grounded in the context-specific constraints, and properly designed based on a realistic theory of change, the support of partners on the ground to ensure compliance is also key.

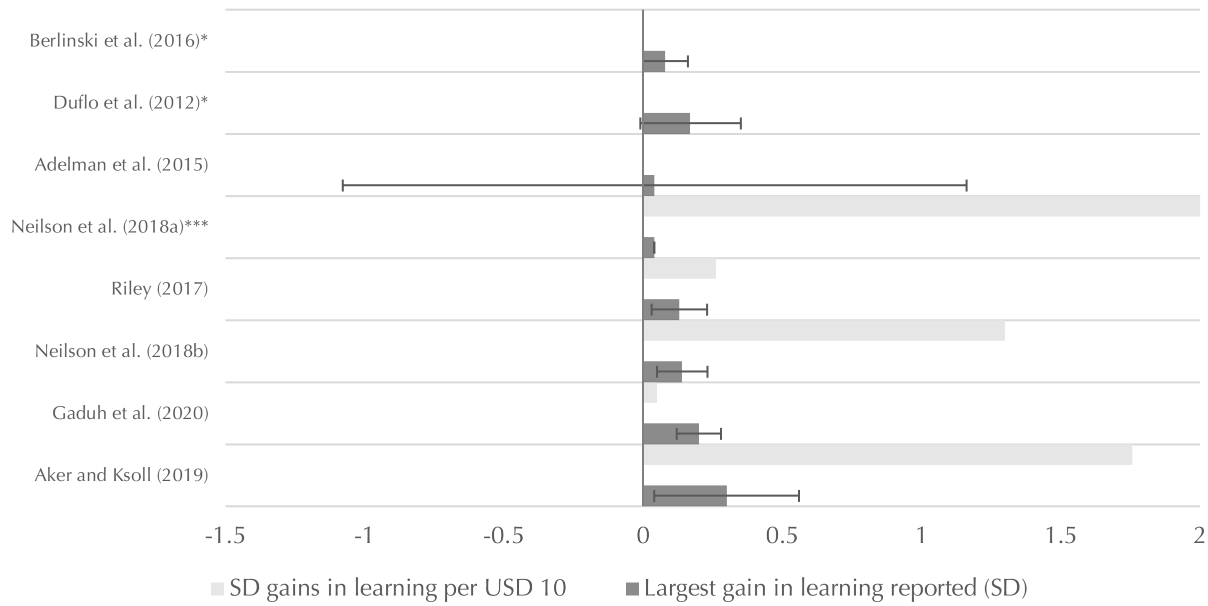
In terms of interventions that are intended to provide information as opposed to increase accountability, there are several examples of interventions that were effective, and highly cost-effective. At the parent-level, Berlinski et al. (2016) evaluate a program which consisted of high-frequency texting campaigns for parents in Chile, during which they were informed about their children’s performance, attendance, and behavior. The study finds large effects in test scores and attendance after only four months of the intervention, highlighting the crucial role that solving information asymmetries between parents and students can play in keeping students accountable for their school performance. At the student-level, interventions like Neilson et al. (2018a, 2018b) provided students with information on the returns to education through contextually-sensitive videos and infographics, which also had significant effects on the students’ performance and aspirations. Similarly, Riley (2017) leverages “role model” effects through the showing of the movie “Queen of Katwe” to Ugandan students, with positive results in the short term. At the teacher- and school officer-level, interventions like Dustan et al. (2019), and Vakis and Farfan

(2018) also proved successful by sending these stakeholders SMS messages with things like reminders about deadlines, framed using insights from behavioral science such as the inclusion of the recipient's name in each text. Although most of these informational campaigns have effects on the smaller side (i.e. less than 0.10 SD), it is also noteworthy how inexpensive and scalable these interventions really are. Once a system that automates the sending of messages through platforms like WhatsApp or even SMS is in place, the marginal cost of adding new users is extremely low.

Given the smaller size of the effects of information campaigns, this type of intervention does not emerge as a promising lead reformer of educational systems in developing countries. However, their high cost-effectiveness and potential for scalability emphasizes the need to complement other core educational policies with this kind of intervention, which bridges gaps in knowledge and cognitive bandwidth. An important feature shared by all these studies was that the information provided is actionable, relevant for the specific context, and concrete-enough to not overwhelm the recipient, therefore making the translation between new information and improved educational practices easier. Similarly, interventions aimed at improving accountability around the stakeholders of education seem promising, albeit more sensitive to challenges with implementation, monitoring, and scalability. If implemented correctly, these can achieve large gains in academic outcomes such as in Duflo et al. (2012), and very high cost-benefits ratios such as in the case of Aker and Ksoll (2019). However, the support of local partners to design, deploy, and incentivize the take-up of the intervention is crucial, as best exemplified by Adelman et al. (2015).

In all, the extant evidence suggests that properly designed and implemented technology can shape the behavior of education stakeholders in a way that can be scalable and cost-effective, and is indeed a promising area for future research. Instead of a unified global agenda, this particular area calls for in-depth knowledge of contexts, and local constraints which may be alleviated through technology-led interventions. Having said this, issues such as the use of technology to aid parents directly support their children's studies such as in Doss et al. (2018), and the potential for technological channels to inform students about opportunities and deadlines to further their education such as in Castleman and Page (2015) remain fairly unexplored in developing contexts.

Figure 2: comparison between gains in learning reported and cost-benefit ratio for interventions in "Technology-enabled behavioral interventions" category



Notes: in order to display the full potential of each intervention, "learning gain" coded as the largest gain in any field of learning, whether it is an academic subject like "math" or a less established area like "computer skills". Studies denoted with three stars (***) had such a high cost-effectiveness ratio that the bar was recoded as a 2 to ease the visual interpretation of the other studies. In the case of Neilson et al. (2018a), the largest gain in learning corresponds to 6.7 SD per USD 10. The "SD gains in learning per USD 10" corresponds to the largest effect in any field of learning, divided by the per-pupil spending in USD, divided by 10. Studies for which authors did not report enough information to standardize gains into SD units are not in this plot. Studies denoted with a star (*) did not report enough cost information to obtain a per-pupil estimate, and hence a cost-benefit ratio. Studies without confidence intervals did not report standard errors in the results. Studies are sorted by whether they do not have cost information first, and then by the "Largest gain in learning reported" bar.

c. Improvements to instruction

The "improvements to instruction" category includes all interventions aimed at addressing any of the constraints that make the quality of teacher instruction not the best that it could be at boosting learning outcomes. Within this category, I have identified three main sub-themes: remote instruction, shaping of classroom instruction, and remote engagement with teachers and parents. As such, the first sub-theme deals with connecting students with knowledgeable, engaging, and curriculum-specific remote instruction. Figure A8 shows that it is very common for teachers in developing countries to not master the content that they are expected to teach. Therefore, even if other constraints like an appropriately targeted and paced curriculum or the high pupil-teacher ratios were relieved, it is unlikely that students would learn much if teachers do not have a deep knowledge of what they are supposed to teach. The issues around teacher mastery of the content run deep within the structural setup of educational systems. Factors such as teacher recruitment and deployment in "undesirable" areas such as remote regions or places of extreme deprivation, lack of regional teacher formation centers in the more rural areas, and lack of incentives for professional development may also play a crucial role in this issue[32]. Therefore, a substantial portion of the literature has focused on using technology to bring education to the most remote places, or schools with generally weak-performing teachers.

Johnston and Ksoll (2017), Naik et al. (2016), and Bianchi et al. (2019) evaluate the impact of remote instruction via satellite in Ghana, India, and China respectively. As an illustration of this type of intervention, Johnston and Ksoll (2017) evaluated the broadcasting of live instruction via satellite to rural primary school students, from a recording studio in Accra where qualified teachers would lead the lessons for students in grades 2-4. All three studies find significant learning gains in at least one subject. Furthermore, their cost-effectiveness is promising, especially since most of the costs are fixed, making the marginal costs of additional students or even classes very low. Among these, Naik et al. (2016) is particularly remarkable due to their explicit decision to study a program at-scale, reaching almost 2,000 public and private schools across the entire state of Kartanaka. By implementing this program at scale, the authors lower their per-pupil costs to less than USD 2 per year, without necessarily compromising the strong learning gains in the order of 0.10 SD-0.40 SD (depending on the subject). The model of remote instruction was not exclusively tested for live lessons, but also through audio and video recordings. Studies like Beg et al. (2019), Näslund-Hadley et al. (2014), and Wennersten et al. (2015), in Pakistan, Paraguay, and India respectively, also studied the effect of delivering content that complements classroom instruction through pre-recorded content. For example, Beg et al. (2019) delivered expert content through pre-recorded content tailored to the local context, which replaced regular class time and gave teachers tools to review the content of the videos through multiple-choice testing. Näslund-Hadley et al. (2014) was also an intervention with a high degree of local adaptation, as the content of the recordings followed the national math curriculum for preschool, and was taught bilingually in Spanish and Guaraní to mimic the teaching conditions of Paraguayan schools. Along the same lines of pre-recorded videos, the different evaluations of local adaptations of Sesame Street for different contexts (Borzekowski (2018) in Tanzania, Borzekowski and Henry (2010) in Indonesia, Borzekowski et al. (2019a) in Rwanda, Borzekowski et al. (2019b) in India) also all had positive effects on early numeracy and literacy skills of young children. Finally, Angrist et al. (2020) explore the effectiveness of phone-based instruction in Botswana during COVID-induced school closures, showing benefits in the order of 0.3 SD for a weekly 15-20 minute call. This model seems

[32] For instance, see Huang et al. (2020) for a clear illustration of serious teacher recruitment issues in Indonesia.

especially appealing during emergencies as calls must be tailored to each student's level, and feature phones are highly prevalent in the developing world.

The second sub-theme within this category was the complementing and shaping of teacher instruction, as opposed to substitution. The most fitting example is Böhmer et al. (2014), which studied an after-school computer-assisted program in Cape Town focusing on each student's particular weaknesses in math, and giving students agency to pick whichever topics they wanted to work on. This program proved effective at improving math knowledge, and interestingly, it raised foundational math knowledge more than it improved the grade-specific knowledge of students. In other words, by fully customizing the study program to each student's particular weaknesses, this program filled in content gaps that regular instruction might not have remedied, as foundational math skills were already assumed in the grade students were. Three other interesting studies in this sub-category, which also intersect with the broader subsection of "Access to technology" are Berlinski and Busso (2017), Lehrer et al. (2019) and Blimpo et al. (2020). The latter two studies find that providing technology which also enabled improved instructional methods through features like lesson scripts (as in Blimpo et al., 2020) led to better test scores in Senegal and The Gambia respectively. An interesting feature of Blimpo et al. (2020) is that it consists of a very comprehensive treatment that improves access to technologies for teachers and students, but also supports targeted at improving instruction and student engagement. Therefore, the researchers cannot untangle the individual effects of each part of the treatment, and cannot ensure that all the gains were truly due to the portions targeted at improving actual classroom instruction.

Perhaps the most interesting case in this category, and certainly the exception in terms of effect size and direction, is Berlinski and Busso (2017). This study used 85 high schools across Costa Rica, targeting the seventh grade math curriculum, and providing a new non-EdTech instructional approach to encourage "active learning" in geometry. On top of this basic treatment, the study also tested the overlapping provision of different technologies such as interactive whiteboards, computer labs, and computers for each student across the different experimental arms. The authors find that no treatment arm had positive effects on learning, the intervention that simply had an instructional change to encourage active learning had *negative* effects in the order of -0.17 SD, and the treatment with active learning plus technology has negative effects in the order of -0.25 SD. Unlike in Adelman et al. (2015) in the previous section, the teacher take-up for this intervention was high, and it was implemented as expected. Instead, the authors attribute the negative results to worsened interactions between the teachers and their students, as evidenced by the negative effects on student discipline, and the teachers' feelings of worst control over the classroom management. This study acts as a cautionary tale warning against sudden instructional and curricular changes, particularly when these come with significant technological adjustments in the classroom.

The third sub-theme in this category is remote coaching and meetings, as best exemplified by Kotze et al. (2019) and its three-year follow up by Cilliers et al. (2020), and Wolf et al. (2018), in South Africa and Ghana respectively. These programs leverage technology to connect remotely with teachers and parents. In the case of Wolf et al. (2018), the authors integrate technology as a component in a broader treatment arm which intended to get parents more involved with the intervention. While the teacher training intervention was *less* effective when parents were involved, the bundled treatment does not allow the researchers to tease apart the effect of purely online meetings. On the other hand, Kotze et al. (2019) explicitly tested a virtual teacher training module versus an on-site training, in light of questions regarding the scalability of on-site coaching for teachers. The authors find that they both had similar positive effects, but the virtual training was slightly cheaper, and signified a less logistically-challenging task to scale than on-site coaching, in spite of the three year follow up of the study (Cilliers et al., 2020) showing diminishing returns to virtual coaching in the longer term. Finally, one important consideration for the rollout of virtual training is that teachers had to be provided with tablets, which even if it is cheaper than on-site training, may still require access to electricity.

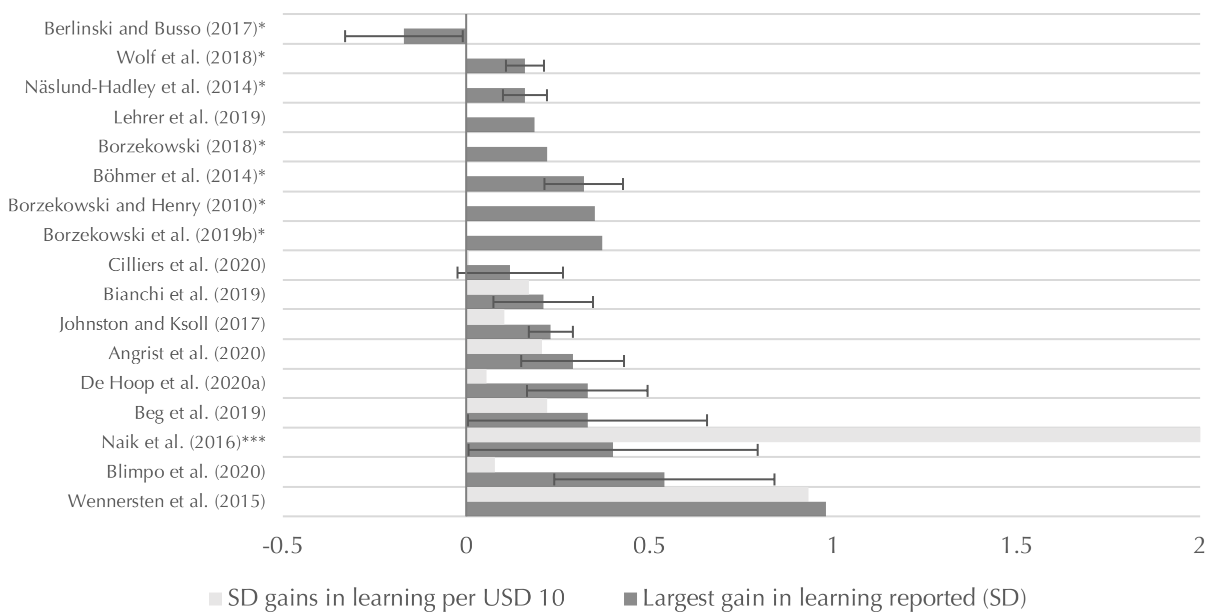
In all, the current evidence points to the fact that the "Improvements to instruction" category is a very promising area for the use of EdTech in developing countries. In fact, the median effect size among all studies reviewed was 0.28 SD, and the 75^{th} percentile was 0.38 SD. Throughout most of the interventions reviewed here, the proper identification of contextual binding constraints when it comes to instruction seems to be a common thread. The design of the intervention around the issue at hand was key at improving learning levels, whether this constraint was teacher knowledge or effectiveness like in the case of Beg et al. (2019), or the scalability of teacher coaching systems, such as in Kotze et al. (2019) and Cilliers et al. (2020). A large portion of the studies focused on a model of partially replacing or supplementing some classroom instruction through technological tools like live broadcasted lessons, pre-recorded videos, T.V.

shows, and audio recordings. This model of EdTech delivery acknowledges the diminishing returns from teacher instruction in contexts where teachers may not fully master the content they are expected to teach, or cannot deliver said content to the full range of achievement levels within their classrooms.

Having already discussed the promising role for this type of interventions, it is important to also mention that none of the papers included here speak to whether EdTech can fully replace classroom instruction. This is a crucial question, especially if schools are not only thought of places to build academic skills, but also a place of socioemotional and psychological development. Furthermore, given the key role of locally-identified constraints in the effectiveness of this type of intervention, none of the papers reviewed seem to suggest that *all* EdTech interventions which address shortfalls in instruction through complementation or replacement of time work. In fact, Berlinski and Busso (2017) serves as a stark reminder of an intervention which had negative effects, and were only aggravated by the use of technology. While the current literature empirically explores cases of EdTech ameliorating learning through improvements in instruction, there still needs to be more research on what areas of the classroom experience are riper for this type of intervention. In other words, if there is at least one study with negative effects, and other studies with different magnitudes for their positive effects, there is a possibility that EdTech can play different roles when it comes to substituting or complementing instruction. Therefore, future areas of research could explore whether EdTech is more effective at replacing actual instruction or at reinforcing instruction through tailored exercises after an actual teacher lecture. Similarly, future research could inform what teacher and school characteristics are more predictive of effective classroom instruction replacement by EdTech components.

EdTech can also be leveraged to incorporate other changes to instructional methods. For instance, scripts which the provide scaffolded lesson plans to teachers have been a part of successful interventions in several developing countries (Piper et al., 2018). Although scripts do not necessarily have to be delivered through a technological device, education providers such as Bridge International Academies already leverage handheld devices connected to the internet to routinely deliver structured lessons at-scale to all of their teachers across several developing countries. While scripts have been part of promising interventions that have raised literacy outcomes for children, no impact evaluations of purely teacher scripts were located for this review, much less as delivered by electronic devices. Similarly, there are no publicly available impact evaluations of different features in teacher scripts and how these affect the quality of instruction, in spite of the valuable descriptive analyses in Piper et al. (2018) and Piper and Evans (2020).

Figure 3: comparison between gains in learning reported and cost-benefit ratio for interventions in "Improvements to instruction" category



Notes: *in order to display the full potential of each intervention, "learning gain" coded as the largest gain in any field of learning, whether it is an academic subject like "math" or a less established area like "computer skills". Studies denoted with three stars (***) had such a high cost-effectiveness ratio that the bar was recoded as a 2 to ease the visual interpretation of the other studies. In the case of Naik et al. (2016), the largest gain in learning corresponds to 2.4 SD per USD 10. The "SD gains in learning per USD 10" corresponds to the largest effect in any field of learning, divided by the per-pupil spending in USD, divided by 10. Studies for which authors did not report enough information to standardize gains into SD units are not in this plot. Studies denoted with a star (*) did not report enough cost information to obtain a per-pupil estimate, and hence a cost-benefit ratio. Studies without confidence intervals did not report standard errors in the results. Studies are sorted by whether they do not have cost information first, and then by the "Largest gain in learning reported" bar.*

d. <u>Self-led learning</u>

The success and cost-effectiveness from the evaluation of the MindSpark software in Muralidharan et al. (2019) sparked great interest in technological interventions which allow students to learn at their own pace, and at their own level. EdTech interventions that enable students to learn at a fitting pace with minimal external support seem particularly enticing, especially in contexts where regular classroom instruction may not be as effective, and there are important resource constraints in terms of teacher and tutor time to ensure that all children make similar progress. Furthermore, interventions that target "self-led learning" have been one of the main areas of EdTech research, accounting for a third of all core studies identified by this review, and dating back to at least 2003 (Rosas et al. in Chile). While it is difficult to draw a sharp distinction between "self-led learning" and "improvements to instruction", the general spirit of "self-led learning" is precisely interventions that students can do mostly on their own, and do not necessarily intend to improve the overall classroom instruction as a mechanism to achieve higher learning, but rather to deliver content directly to students. Similarly, unlike in the "access to technology" category, most of the interventions in this category did not provide students with the hardware or the devices to engage with the intervention and instead, most self-led activities were software-oriented. While it would be possible to implement an intervention which merges "access to technology" and "self-led learning" at an individual level (e.g. through the provision of a handheld device with appropriate self-led software installed), most of the interventions in this category leveraged technology at the school-level. By targeting communal sharing of the hardware to implement self-led interventions, the marginal costs spread out further than initiatives like OLPC, as it allows several students to use the same hardware during a school year, and then for several cohorts to keep using until it fully depreciates.

Interestingly, the majority of all studies in this section had at least one treatment arm with positive effects on learning, as Figure 4 shows. In fact, the median effect size in this category is 0.29 SD, and the 75th percentile is 0.46 SD. Therefore, the bulk of the evidence in this section does not revolve around whether there is a model of self-led learning which works, but rather around how different design features of self-led learning interventions moderate the effects that these have on learning outcomes. Two important exception of this are Büchel et al. (2020) and Ma et al. (2020), which instead of testing a different feature of an EdTech intervention, evaluate an EdTech intervention in relation to a comparable "pencil-and-paper" treatment. In the case of Ma et al. (2020), the authors highlight that EdTech interventions, particularly those in this category, tend to happen after school. Therefore, there is a question about whether any learning gains observed are due to the EdTech portion of the intervention, or rather due to the additional practice time. The authors find that for their particular treatment, the EdTech treatment branch is no more effective than the non-EdTech arm, suggesting that part of the success of interventions in this category may be because it offers students additional practice time. On the other hand, the authors of Büchel et al. (2020) test whether students assigned to computer-assisted learning (CAL) fare better than those in a traditional teaching environment during a weekly, 90-minute intervention, finding that CAL is indeed more effective than traditional teaching in their context. The contrast between these two interventions may lie in the contextual counterfactual for each. While the Ma et al. (2020)

study was conducted in China, the Büchel et al. (2020) study was conducted in El Salvador, a country with a lower development level, and weaker state capacity that may translate into a poorer traditional classroom experience. Hence, this difference highlights the importance of clearly understanding the contextual constraint that an EdTech product would address, and the resources that it would be displacing if implemented. Having said this, there may be features inherent to self-led EdTech interventions that can still make EdTech desirable over non-EdTech interventions, or business-as-usual teaching. For instance, EdTech software has the capacity to hold a very large number of questions, with a wide range of difficulty, and with minimum setup and external support, allowing for greater scalability and extended exposure to each intervention.

One of the first design features that the literature touches upon is the difference between "computer-assisted instruction" (CAI) and "computer-assisted learning" (CAL). Although some authors use the terms interchangeably, the clearest distinction is drawn by Bai et al. (2016). This study defines CAL as not necessarily integrated into the teachers' instruction and curriculum, whereas CAI is. In fact, Bai et al. (2016) test this distinction explicitly in their experimental design, by comparing CAI and CAL treatment arms to a pure control group, finding suggestive evidence that CAI was more effective than CAL at raising English test scores. More broadly, other papers tested one or the other model without explicitly defining their intervention as CAL or CAI. Linden (2008) is an informative paper in this regard, particularly as it also studies the properties of EdTech as supplements or complements to math instruction in Gujarat, India. Linden (2008) compares a computer-led intervention implemented as an in-school program ("substitute" of in-class instruction), or out-of-school addition ("complement" of in-class instruction) on second and third graders. The author finds that the intervention had negative effects as a supplement of instruction, but the intervention had positive effects in the order of 0.3 SD when it was used as a complement to reinforce instruction, effectively being used as CAI. Other interventions such as He et al. (2008) were leaning more towards the CAL side, as it was focused on self-exploration of topics within a specialized device, also yielding positive effects. In this sense, the difference between these two approaches is not necessarily along the margin of whether one is strictly better than the other, but which one is better suited for the task at hand. Work such as Bai et al. (2008), Lai et al. (2013, 2015, 2016) or Mo et al. (2014a, 2014b) highlights the virtue of CAI to act as a complement to in-class instruction and content, while work such as Linden (2008), Bettinger et al. (2020), Carrillo et al. (2011), Chong et al. (2020), or Rosas et al. (2002) displays the potential of CAL to reinforce concepts that do not precisely mimic the students' curriculum at any specific point in time. For instance, Chong et al. (2020) targets sex education for Colombian teenagers, and stands as a valuable example of a case when CAL may be more effective than CAI, especially if the content delivered in class would either be poorly communicated at school or not taught at all.

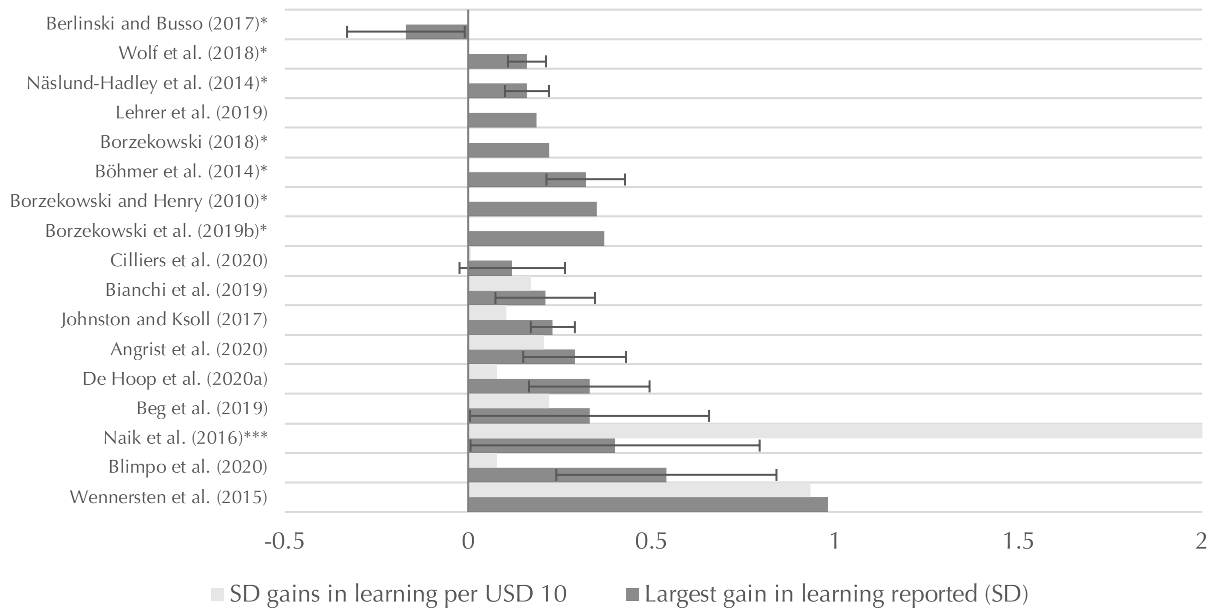
Another important design feature that has captured little research attention across the papers in the set of core studies is the incentives provided to students to engage with EdTech products. Hirshleifer (2016) is the only study included in this review which explicitly evaluates two different incentive schemes. Specifically, the author studies whether rewarding "inputs" or "effort" to engage with an EdTech product is more effective than rewarding "outputs" or the actual score obtained on the EdTech activity. Hirshleifer (2016) finds that for their specific intervention, rewarding inputs is more than twice as effective as rewarding outputs, although both modalities of rewards yield important learning gains. However, this paper only deals with one type of small reward with a maximum value of USD 2.65 per child, and does not test different types of rewards such as social recognition, symbolic gestures of teacher appreciation, or the potential to earn a significantly larger prize. Similarly, work such as Araya et al. (2019) or Rosas et al. (2002) recognize the potential for gamification in driving engagement with an EdTech product. In a qualitative analysis into potential mechanisms for their lack of significant results, De Hoop et al. (2020b) find that some characteristics of their product seemed repetitive, and led to boredom for the students using the software they evaluate. Therefore, including features that touch upon "gamification" to drive engagement with EdTech products could potentially be an even more cost-effective incentive. Still, so far none of the studies included in this review explicitly tests the sole effect of features like gamification on the effectiveness of an EdTech product.

A key component of some EdTech products which has not been evaluated in isolation is the optimal degree of adaptability, i.e. the potential for the product to auto-identify and adjust the level of difficulty to a student's specific achievement level. This particular feature has been a core component of very successful interventions such as Banerjee et al. (2007), Muralidharan et al. (2019), Ito et al. (2019), and Carrillo et al. (2011). Given the wide variation in achievement distributions within classrooms in many developing countries, this feature is one of the most enticing characteristics of EdTech, and it is hard to imagine that it would be anything but beneficial for each student's learning path. Therefore, the key empirical question around adaptability is not whether it works or not, but rather what the optimal degree of adaptability is. This is relevant since there are certainly higher development costs to creating deeper question banks with different difficulty levels, and to the ideation of more sophisticated algorithms to precisely place students within the performance bin that the EdTech product would target. In spite of the potential relevance for policymakers and product developers, no paper in the current set of core studies directly addresses this question in a self-led learning intervention.

The final feature discussed in this review[33] for which little evidence currently exists is the optimal dosage for an intervention. All interventions in this category have different lengths for their study sessions, and different number of weeks during which students were a part of the intervention. However, only Bettinger et al. (2020) explicitly tests the effect of different dosages of an EdTech intervention. The authors find that while the treatment does have positive effects on learning, the full doubling of the dosage does not have statistically different effects from the baseline intervention. This finding agrees with the null correlation found between dosage and effect size across different studies in Sampson et al. (2019). Understanding this relationship is crucial when deciding not only whether EdTech should be a complement or a supplement to education, but also to what degree it should be implemented as either. Furthermore, dosage is an important feature given the nature of *self-led* interventions, where the learner must have some autonomy, and the ability to understand how the product works. An intervention with a long dosage period, but which low-performing students struggle to engage with, is likely to have heterogenous effects across the full distribution of achievement, ultimately benefiting stronger students and widening within-class and within school inequality. In fact, Carrillo et al. (2011) and He et al. (2008) observe that higher-performing students perform better their self-paced EdTech interventions. Therefore, the suitability of the treatment for the specific context, adaptability for different learning levels, and crucially, the right dosage for everyone's needs are pivotal elements to ensure that self-led EdTech interventions can cater and boost educational outcomes for all students.

[33] Note that this is not a comprehensive list of potential features to be studied and/or included in an EdTech product. Sampson et al. (2019) mentions other potential features which an EdTech product could include, such as the inclusion of different components like "explanatory videos", "practice exercises", "problem solutions", "assessments", "quizzes/stories", "simulations", "flash cards", among others.

Figure 4: comparison between gains in learning reported and cost-benefit ratio for interventions in "Self-led learning" category



Notes: in order to display the full potential of each intervention, "learning gain" coded as the largest gain in any field of learning, whether it is an academic subject like "math" or a less established area like "computer skills". Studies denoted with three stars (***) had such a high cost-effectiveness ratio that the bar was recoded as a 2 to ease the visual interpretation of the other studies. In the case of Hirschleifer (2016), the largest gain in learning corresponds to 2.2 SD per USD 10. The "SD gains in learning per USD 10" corresponds to the largest effect in any field of learning, divided by the per-pupil spending in USD, divided by 10. Studies for which authors did not report enough information to standardize gains into SD units are not in this plot. Studies denoted with a star (*) did not report enough cost information to obtain a per-pupil estimate, and hence a cost-benefit ratio. Studies without confidence intervals did not report standard errors in the results. Studies are sorted by whether they do not have cost information first, and then by the "Largest gain in learning reported" bar.

VI. Lessons learned and frontiers of the current evidence

The current review provides a comprehensive compilation of rigorous EdTech interventions in developing countries. By thematically grouping all 81 core studies, broader lessons can be drawn for future research and implementation of EdTech interventions, as synthesized in Table 5. Among the four categories, the most promising areas in raising learning outcomes were "improvements to instruction" and "self-led learning." The overall success of these two areas rested on the customization of the EdTech solution to the constraint at hand. The studies included in "improvements to instruction" addressed more systematic constraints such as weak teacher quality in certain remote areas or teaching coaching through scalable, virtual means. The "self-led learning" studies focused more on a direct link connecting students to learning through technology like apps or educational software. At the same time, "technology-enabled behavioral interventions" also seems to be particularly effective at solving problems of informational-asymmetries, accountability and enforcement of duties, while also being particularly cost-effective and prone to scalability. The studies under "access to technology" did not show a pattern of raising learning, only students' acquaintance with technology. However, interventions that facilitate access to technology are a first and necessary step to implement other EdTech solutions like educational software, especially in many remote and deprived areas. Most importantly, there is a need for researchers and policymakers to move away from a dogmatic adherence to one of the four areas, and to embrace the fact that all four areas can act as mutually complementary in addressing deficiencies within educational systems.

Table 5: summary of EdTech interventions in developing countries by thematic area

	Access to technology	Technology-enabled behavioral interventions	Improvements to instruction	Self-led learning
Intended policy targets	Low penetration of technologies able to host educational features, low familiarity with digital skills.	Informational barriers, behavioral inconsistencies, lack of accountability, alignment of incentives.	Gaps in teacher knowledge, difficulties to recruit teachers in remote areas, scalability of student and teacher training programs.	Reinforcement of material and practice problems, addressing student-specific gaps in skills, adjusting the pace and level of instruction.
Effectiveness	Very low for academic learning, medium for increases in familiarity with digital tools.	Low to medium-sized effects in learning.	Consistently medium to large effects in learning.	Among the software evaluated, consistently medium to large effects in learning.
Cost-effectiveness	Extremely low. Poor effectiveness coupled with high marginal costs. As a result, expensive to scale.	Very high, particularly due to the very low marginal costs of most interventions. Very high potential for scalability.	High, as fixed costs of product development tend to be higher than marginal costs.	High, as interventions are often implemented in community- or school-level computer labs so the same hardware/software can reach many students.
Best uses	Increase familiarity with technology, or as a platform to implement other types of EdTech interventions.	Improve enforcement of policies, provide information at scale.	Deliver high-quality education to areas where this is a serious constraint.	Complement classroom instruction, reinforce lessons, fill in content gaps.
Potential pitfalls and challenges	Leakage and misuse of equipment, crowding out of time better spent in other educational activities.	Interventions require particularly deep contextual knowledge about behaviors that can be shaped through relatively low-touch interventions.	A sudden change in technology that does not directly address a pressing problem may hinder instruction and lead to negative effects in learning.	Software needs to be developed for more contexts, languages, and subjects. Reliance on self-guidance may benefit high achievers more, increasing within-class inequality.
Examples of interventions	One-laptop-per-child (OLPC) (Barrera-Osorio and Linden, 2009; Cristia et al., 2017), provision of handheld devices (Habyarimana and Sabarwal, 2018; Mensch and Haberland, 2018)	Keeping parents up to date on student performance and attendance via SMS (Berlinski et al., 2016). Monitoring teacher attendance through cameras linked to pay incentives (Gaduh et al., 2020)	Broadcasting of live instruction remotely (Johnston and Ksoll, 2017). Pre-recorded video and audio lessons to supplement classroom instruction (Beg et al., 2019; Näslund-Hadley et al., 2014)	Software (typically self-adaptive) to practice language and math skills (Muralidharan et al., 2019; Linden, 2008; Carrillo et al.; 2011; Araya et al., 2019). Online classes (Chong et al., 2020).

Another important lesson that emerged from the four thematic areas is the importance for an EdTech intervention to be thoughtfully designed around a carefully identified contextual issue. To illustrate this point, one can look at the way in which Beg et al. (2019) identify clear contextual constraints: unavailability of qualified teachers and teacher absenteeism respectively; they hypothesize about appropriate and scalable technological approaches to address these issues with contextually-grounded theories of change through the provision of short videos with academic content in math and science, which led to large and cost-effective gains in learning and some evidence for increased teacher effort. The implementation of this program was during school time, and through the local government. This intervention stands in sharp contrast to Angrist and Lavy (2002) or even the OLPC interventions, which attempt to address a more nebulous issue of access to computers without a clear theoretical, causal path between owning a computer to improved school performance. In the extreme case of Angrist and Lavy (2002), a well-intentioned and expensive intervention ended up even yielding negative results in learning.

The quality of implementation and take-up from relevant stakeholders also stand as pivotal components to understanding the success or failure of an intervention. However, quality of implementation does not seem to replace a well-thought out design. In other words, while quality of implementation could make or break a project that may be indeed appropriate to address certain issues if properly implemented, such as in Adelman et al. (2015), a successful implementation and take-up does not guarantee gains in educational outcomes. As an illustration of this point, Berlinski and Busso (2017) report high take-up of their treatment, and no issues with implementation are reported. However, the intervention also led to negative effects, being worsened by the inclusion of technology into the change in pedagogical approach. While an initial reaction to this major point about quality of implementation may be to motivate implementers of the study to exert exceptional effort and resources to ensure that the intervention goes precisely as planned, the end goal for most of these interventions is to test whether they have a potential for scalability. In many cases, the difficulty of maintaining a high level of quality in the implementation phase tends to get larger with the size of the intervention. Therefore, a lesson that emerges from this review, and from other work like Niehaus and Muralidharan (2016) for that matter, is to give preference to intervention designs with relatively few touchpoints between the delivery of treatment and the target population, so that if and when the intervention is scaled, it can adhere to similar implementation standards as in the pilot.

Relatedly, the question of scalability also emerges as an important issue when it comes to EdTech interventions. For instance, an interesting feature for EdTech interventions is the interplay between fixed and marginal costs[34]. Depending on the type of intervention, there could be serious trade-offs between the two types of costs that could significantly affect scalability and economies of scale in expanding treatment to other individuals. Two opposite examples are the OLPC studies (Barrera-Osorio and Linden, 2009; Beuermann et al., 2015; Cristia et al., 2010, 2017; de Melo et al., 2014; Meza-Cordero, 2017) versus the "Sesame Street" studies (Borzekowski, 2018; Borzekowski, 2010; Borzekowski et al., 2019a; Borzekowski et al., 2019b). The nature of OLPC policies is that the cost of adding an additional child is exactly the cost of a laptop. There may be some economies of scale through lower prices when buying computers in bulk, but the marginal cost is still considerably higher than any fixed costs per student associated with running the program. Contrarily, the cost of "Sesame Street"-type interventions is mostly focused around the fixed-costs of developing, producing, and distributing the T.V. episodes. However, the marginal cost of another student watching the show is effectively zero. Unsurprisingly, most of the studies reviewed here lie somewhere in between these two extremes, and their position along this spectrum also depends heavily on the area of the review. For instance, interventions within the "access to technology" category tend to skew towards higher marginal costs, and interventions within the "improvements to instruction" tend to skew towards higher fixed costs. This distinction is crucial to welfare analyses of EdTech interventions, as interventions with low marginal costs and positive effects, as small as they may be, stand to achieve Pareto improvements by enrolling more children, while interventions with high marginal costs must consider more carefully whether the marginal benefit to the infra-marginal student will indeed justify the relatively higher costs.

Another potential consideration for the scalability of EdTech products is the trade-off between the economies of scale of product development, and the tailoring of a product to the local context. In other words, the larger the market an intervention intends to target, the more costly the tailoring of the intervention would be. For instance, an EdTech solution focusing on early language development in a country with many regional languages would either need to develop a different version for each regional language, or focus on the main national and/or colonial language, which may also have equity implications. A similar pattern occurs across different grades: while most early curricula in most countries focuses, in one way or the other, on the development of foundational literacy and numeracy skills, the contents of curricula grow increasingly different

[34] For an excellent review of the advantages, and necessary conditions for the successful scalability of interventions in developing countries, see Niehaus and Muralidharan, 2016.

across countries with grade progression. Therefore, an app focusing on early skills may have a larger potential market than one focusing on a niche curricular feature, such as pre-colonial Nigerian history, which may be present in Nigeria's curriculum but not Ghana.

Given the inherent limitations, costs, and barriers to entry that EdTech interventions may face, it is also important to note that from the core set of studies, it is not clear whether EdTech interventions *always* achieve higher learning gains and are *always* more cost-effective, compared to other non-EdTech interventions in developing countries. In this sense, the question that policymakers and researchers face when evaluating an EdTech intervention should not be whether this technological approach could address a problem in the educational system, but rather whether it would be the most effective and cost-effective way to do so. Indeed, there are examples of non-EdTech interventions in developing countries that have been equally as successful at raising learning standards as the most promising EdTech solutions, such as "Teach at the Right Level" (Banerjee et al., 2016) or the combination of other fruitful approaches such as scripting and after-school remediation lessons (Eble et al., 2019). Besides the cost and ease of implementation and scalability, the decision to implement an EdTech intervention versus an equally well-designed non-EdTech solution should come down to whether the intervention could benefit from the comparative advantages offered by EdTech, such as the potential for high levels of customization of practice exercises or remote engagement.

Among the set of broader questions that remain on the frontier of EdTech research are those involving "general equilibrium" effects after the rollout of an EdTech intervention. Very little is known about the system-level, medium- and long-term effects on teacher attitudes, effort, and behavior following an EdTech intervention. One can imagine a context where teachers quickly adapt the technology to their daily routine and set of tools, as it becomes an integral part of education. Conversely, there could also be a scenario in which the take-up of technology only happens during a brief period of excitement or monitoring, and the use is then gradually discontinued. Similarly, one can imagine teachers feeling more motivated about new technology lifting some of their instructional burden and hence putting more effort into the time that they actually teach, or on the contrary, teachers relying on EdTech as a substitute of instruction to maintain or increase their absenteeism rates. Questions of this nature can be asked at the school-level and even at the system-level, where it is unclear whether EdTech can crowd out resources of other important educational inputs, or will instead boost the effectiveness of other complementary investments. Similarly, little is known about the susceptibility of EdTech interventions to political and investment cycles[35]. Conditional on finding a set of interventions that raise educational outcomes in a specific context, the continuity of these programs by future education leaders and policymakers is just as crucial as the finding that the intervention is an effective one.

The breadth in the EdTech literature, in terms of type of intervention and context, is greater than the current depth of it, both in terms of replication of studies in different contexts, and multiple angles to similar research questions. As EdTech keeps growing throughout different developing countries, and policymakers face more options to address the particular challenges in their respective contexts, the body of knowledge in various aspects of when, where, and for whom EdTech interventions work must also grow. Addressing critical questions of scalability, external translation of results, preparedness for EdTech interventions within and between countries, and the particular shortcomings of educational systems in developing countries where EdTech can be most effective will be of paramount importance to keep up with an evidence-based agenda in pursuit of improved educational and welfare outcomes for people in the developing world.

[35] For informative case studies on how South Korea, Estonia, and Uruguay have integrated ICT into their educational system at-scale, see Díaz et al., 2020.

VII. References

ASER. (2014). ASER India data-Household data [Data file]. http://www.asercentre.org/education/data/india/statistics/level/p/66.html

ASER. (2016). ASER Pakistana data-Household data [Data file]. http://aserpakistan.org

Banerjee, A. V., Banerji, R., Berry, J., Kannan, H., Mukerji, S., & Walton, M. (2016). Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of "Teaching at the Right Level" in India. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2846971

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., & Walton, M. (2017). From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. Journal of Economic Perspectives, 31(4), 73–102. https://doi.org/10.1257/jep.31.4.73

Berry, J. (2015). Child Control in Education Decisions: An Evaluation of Targeted Incentives to Learn in India. Journal of Human Resources, 50(4), 1051–1080. https://doi.org/10.3368/jhr.50.4.1051

Bulman, G., & Fairlie, R. W. (2016). Technology and Education. In Handbook of the Economics of Education (Vol. 5, pp. 239–280). Elsevier. https://doi.org/10.1016/B978-0-444-63459-7.00005-1

Castleman, B. L., Page, L. C. (2015). Summer nudging: can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? Journal of Economic Behavior and Organization. 115, 144-160.

Chetty, R., Friedman, J. N., Rockoff, J.E. (2014). Measuring the impacts of teachers, I: Evaluating bias in teacher value-added estimates American Economic Review. 104(9), 2593-2632. DOI: 10.1257/aer.104.9.2593

Chinn, M. D., & Fairlie, R. W. (2010). ICT Use in the Developing World: An Analysis of Differences in Computer and Internet Penetration. Review of International Economics, 18(1), 153–167. https://doi.org/10.1111/j.1467-9396.2009.00861.x

Crawfurd, L., (2020, May 18). Why the COVID Crisis Is Not Edtech's Moment in Africa. [Blog post]. Center for Global Development. https://www.cgdev.org/blog/why-covid-crisis-not-edtechs-moment-africa

De Ree, J., Muralidharan, K., Pradhan, M., & Rogers, H. (2018). Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia. The Quarterly Journal of Economics, Volume 133, Issue 2, May 2018, Pages 993–1039, https://doi.org/10.1093/qje/qjx040

DellaVigna, S., & Linos, E. (2020). RCTs to Scale: Comprehensive Evidence from Two Nudge Units. Working paper. https://eml.berkeley.edu/~sdellavi/wp/NudgeToScale2020-03-20.pdf

Díaz, M. B., Mercedes, M. Lee, C., Zucchetti, A., Olszewski, B., Cobo, C., Viik, L., Kyllönen, M., South, J., Montaldo, M., Ramos, Y. (2020). What Technology Can and Can't Do for Education: A Comparison of 5 Stories of Success. Interamerican Development Bank. http://dx.doi.org/10.18235/0002401

Doss, C., Fahle, E. M., Loeb, S., York, B. N. (2018). More than just a nudge: supporting kindergarten parents with differentiated and personalized text-messages. The Journal of Human Resources. 54(3), 567-603.

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

Eble, A., Frost, C., Camara, A., Bouy, B., Bah, M., Sivaraman, M., Hsieh, J., Jayanty, C., Brady, T., Gawron, P., Boone, P., & Elbourne, D. (2019). How much can we remedy very low learning levels in rural parts of low-income countries? Impact and generalizability of a multi-pronged para-teacher intervention from a cluster-randomized trial in The Gambia. (EdWorkingPaper: 19-157). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/3n1x-rh14

Escueta, M., Nickow, A.J., Oreopoulos, P., Quan, V. (2020). Upgrading Education with Technology: Insights from Experimental Research. Forthcoming in the Journal of Economic Literature.

Evans, D., Mendez Acosta, A., (2020). Education in Africa: What Are We Learning? CGD Working Paper 542. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/education-africa-what-are-we-learning

Evans, D., & Ngatie, M. (2020). School uniforms, short-run participation, and long-run outomes: evidence from Kenya. The World Bank Economic Review, lhaa004, https://doi.org/10.1093/wber/lhaa004

Evans, D., Yuan, F., & Filmer, D. (2020). “Are Teachers in Africa Poorly Paid? Evidence from 15 countries”. Working Paper.

Gallego, F. A., Näslund-Hadley, E., & Alfonso, M. (2019a). Changing Pedagogy to Improve Skills in Preschools: Experimental Evidence from Peru. The World Bank Economic Review, lhz022. https://doi.org/10.1093/wber/lhz022

Gallego, F., Näslund-Hadley, E., & Alfonso, M. (2018). Tailoring Instruction to Improve Mathematics Skills in Preschools: A Randomized Evaluation. Inter-American Development Bank. https://doi.org/10.18235/0001090

Ganimian, A. J., & Murnane, R. J. (2016). Improving Education in Developing Countries: Lessons From Rigorous Impact Evaluations. Review of Educational Research, 86(3), 719–755. https://doi.org/10.3102/0034654315627499

Glewwe, P., & Kremer, M. & Moulin, S. (2009). Many children left behind? Textbooks an test scores in Kenya. American Economic Journal: Applied Economics, 1(1), 112-135. https://doi.org/10.1257/app.1.1.112

Glewwe, P., & Muralidharan, K. (2016). Improving Education Outcomes in Developing Countries. In Handbook of the Economics of Education (Vol. 5, pp. 653–743). Elsevier. https://doi.org/10.1016/B978-0-444-63459-7.00010-5

Global Partnership for Education – GPE. (2019). Improving teaching and learning. Knowledge and Innovation Exchange Discussion Paper. https://www.globalpartnership.org/sites/default/files/document/file/2019-07-kix-improving-teaching-and-learning-final.pdf

Hawkins, R. J., & Freeman, B. (2017). Evoke—Developing skills in youth to solve the World's most complex problems: Randomized impact evaluation findings. World Bank Education, Technology and Innovation : SABER-ICT Technical Paper Series; No. 19. Washington, D.C. : World Bank Group.

Hennessy, S., Harrison, D., & Wamakote, L. (2010). Teacher factors influencing classroom use of ICT in sub-Saharan Africa. Itupale Online Journal of African Studies, 2, 39–54.

Ho, J., & Thukral, H. (2009). Tuned In To Student Success. Assessing the Impact of Interactive Radio Instruction for the Hardest-to-Reach. *Journal of Education for International Development 4:2*

Huang, A.R., Revina, S., Fillaili, R., and Akhmadi. 2020. The Struggle to Recruit Good Teachers in Indonesia: Institutional and Social Dysfunctions. RISE Working Paper Series. 20/041. https://doi.org/10.35489/BSG-RISE-WP_2020/041.

Kerwin, J. T., & Thornton, R. L. (2020). Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures. The Review of Economics and Statistics, 1–45. https://doi.org/10.1162/rest_a_00911

Kozma, R.B., & Surya Vota, W. (2014). ICT in developing countries: policies, implementation, and impact. J.M. Spector et al. (eds.), Handbook of Research on Educational Communications and Technology, 885 DOI 10.1007/978-1-4614-3185-5_72

Kremer, M., Brannen, C., & Glennerster, R. (2013). The Challenge of Education and Learning in the Developing World. Science, 340(6130), 297–300. https://doi.org/10.1126/science.1235350

Lavinas, L., & Veiga, A. (2013). Desafios do modelo brasileiro de inclusão digital pela escola. Cadernos de Pesquisa, 43(149), 542–569. https://doi.org/10.1590/S0100-15742013000200009

Ma, Y., Fairlie, R., Loyalka, P., & Rozelle, S. (2020). Isolating the "Tech" from EdTech: Experimental Evidence on Computer Assisted Learning in China (No. w26953; p. w26953). National Bureau of Economic Research. https://doi.org/10.3386/w26953

Malisa, M., & Missedja, T. Q. (2019). Schooled for servitude: the education of African children in British colonies, 1910-1990. Genealogy, 3(3), 40. https://doi.org/10.3390/genealogy3030040

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., & Rajani, R. (2019a). Inputs, incentives, and complementarities in education: experimental evidence from Tanzania. The Quarterly Journal of Economics, 134(3), 1627-1673. https://doi.org/10.1093/qje/qjz010

Mbiti, I; Muralidharan, K; Romero, M; Schipper, Y; Manda, C; Rajani, R. (2019b). "Replication Data for: 'Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania'", https://doi.org/10.7910/DVN/AKVGQ7, Harvard Dataverse, V2.

Mbiti, I., Romero, M., & Schipper, Y. (2019c). Designing effective teacher performance pay programs: experimental evidence from Tanzania. NBER Working Paper No. 25903.

Mbiti, I., Rodrigue-Segura, D. (2020). “Back to basics: Curriculum Reform and Student Learning in Tanzania”. Unpublished working paper.

McEwan, P. J. (2015). Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments. Review of Educational Research, 85(3), 353–394. https://doi.org/10.3102/0034654314553127

Muralidharan, K., Das, J., Holla, A., & Mohpal, A. (2010) The fiscal cost of weak governance: evidence from teacher absence in India. Journal of Public Economics. 145, 116-135.

Muralidharan, K., & Niehaus, P. (2017). Experimentation at Scale. Journal of Economic Perspectives, 31(4), 103–124. https://doi.org/10.1257/jep.31.4.103

Mwiria, K. (1991). Education for subordination: African education in colonial Kenya. History of Education, 20(3), 261-273. https://doi.org/10.1080/0046760910200306

Piper, B., & Evans, D. (2020, February 27). Guiding Teachers Rather than Scripting Them. [Blog post]. Center for Global Development. https://www.cgdev.org/blog/guiding-teachers-rather-scripting-them

Piper, B., Sitabkhan, Y., Mejía, J., and Betts, K. (2018). Effectiveness of Teachers' Guides in the Global South: Scripting, Learning Outcomes, and Classroom Utilization. RTI Press Publication No. OP-0053-1805. Research Triangle Park, NC: RTI Press. https://doi.org/10.3768/rtipress.2018.op.0053.1805

Pritchett, L., & Beatty, A. (2015). Slow down, you're going too fast: Matching curricula to student skill levels. International Journal of Educational Development, 40, 276–288. https://doi.org/10.1016/j.ijedudev.2014.11.013

Rieble, S., Viteri, A., Vásquez, M., & Arias-Ortiz, E. (2020, May 4). COVID-19 y la gran apuesta del aprendizaje desde casa: ¿funcionará para todos? [Blog post]. Interamerican Development Bank. https://blogs.iadb.org/educacion/es/aprendizajedesdecasa/

Rubagiza, J., Were, E., & Sutherland, R. (2011). Introducing ICT into schools in Rwanda: Educational challenges and opportunities. International Journal of Educational Development, 31(1), 37–43. https://doi.org/10.1016/j.ijedudev.2010.06.004

Sabarwal, S, Evans, D., & Marshak, A. (2014). The permanent input hypothesis: the case of textbooks and (no) student learning in Sierra Leone (English). Policy Research working paper; No. WPS 7021. Washington, D.C.: World Bank Group.

Sampson, R., Johnson, D., Somanchi, A., Barton, H., Ruchika, J., Seth, M., Shotland, M. (2019). Insights from rapid evaluations of EdTech products. The EdTech Lab Series. https://centralsquarefoundation.org/wp-content/uploads/EdTech%20Lab%20Report_November%202019.pdf

SDI – Service Delivery Indicators. (2010-2014). Education data [Data file]. https://www.sdindicators.org/aboutus#where-are-we-now-

Sitabkhan, Y., & Platas, L. M. (2018). Early Mathematics Counts: Promising Instructional Strategies from Low- and Middle-Income Countries. RTI Press. https://doi.org/10.3768/rtipress.2018.op.0055.1807

Stone, R., Hoop, T., Coombes, A., & Nakamura, P. (2020). What works to improve early grade literacy in Latin America and the Caribbean? A systematic review and meta-analysis. Campbell Systematic Reviews, 16(1). https://doi.org/10.1002/cl2.1067

TPC (Todos pelas crianças). (2016). Household data [Data file]. https://palnetwork.org/datasets/

Trucano, M., (2010, May 11). Interactive Radio Instruction: A Successful Permanent Pilot Project? [Blog post]. World Bank Blogs. https://blogs.worldbank.org/edutech/iri

Trucano, M. (2015). Mobile Phones & National Educational Technology Agencies, Sachet Publishing & the Khan Academy: What's happening with educational technology use in developing countries. Excerpts from the World Bank's EduTech Blog (Volume VI). Washington, DC: The World Bank.

Trucano, M. (2016a). SABER-ICT Framework Paper for Policy Analysis: Documenting National Educational Technology Policies Around the World and Their Evolution over Time. World Bank Education, Technology & Innovation: SABER-ICT Technical Paper Series,No. 1;. World Bank, Washington, DC. © World Bank.

Trucano, M. (2016b). Teachers and tablets, mobile devices and "MOOC Times": Innovations in technology use in education in developing countries. Excerpts from the World Bank's EduTech Blog (Volume VII). Washington, DC: The World Bank.

Trucano, M., & Dykes, G. (2017). Building and sustaining national educational agencies: Lessons, models and case studies from around the world. Washington, DC: The World Bank.

Twaweza. (2015). Uwezo data-Household data [Data file]. https://www.twaweza.org/go/uwezodatasets

Waddington, H., Masset, E., & Jimenez, E. (2018). What have we learned after ten years of systematic reviews in international development? Journal of Development Effectiveness, 10(1), 1–16. https://doi.org/10.1080/19439342.2018.1441166

World Bank. (2017). World Development Report 2018: Learning to Realize Education's Promise. The World Bank. https://doi.org/10.1596/978-1-4648-1096-1

VIII. Core studies

Abrami, P. C., Wade, C. A., Lysenko, L., Marsh, J., and Gioko, A. (2016). Using educational technology to develop early literacy skills in Sub-Saharan Africa. Education and Information Technologies, 21(4), 945–964. https://doi.org/10.1007/s10639-014-9362-4

Adelman, M., Blimpo, M. P., Evans, D., Simbou, A., & Yarrow, N. (2015). "Can Information Technology Improve School Effectiveness in Haiti? Evidence from a Field Experiment". Working Paper, World Bank, Washington, DC.

Aker, J. C., & Ksoll, C. (2019). Call me educated: Evidence from a mobile phone experiment in Niger. Economics of Education Review, 72, 239–257. https://doi.org/10.1016/j.econedurev.2019.05.001

Angrist, J., & Lavy, V. (2002). New evidence on classroom computers and pupil learning. The Economic Journal, 112(482), 735–765.

Angrist, N., Bergman, P., Brewster, C., & Matsheng, M. (2020). Stemming Learning Loss During the Pandemic: A Rapid Randomized Trial of a Low-Tech Intervention in Botswana. Working paper. Available at SSRN: https://ssrn.com/abstract=3663098 or http://dx.doi.org/10.2139/ssrn.3663098

Araya, R., Arias Ortiz, E., Bottan, N. L., & Cristia, J. P. (2019). Does Gamification in Education Work?: Experimental Evidence from Chile. Inter-American Development Bank. https://doi.org/10.18235/0001777

Bai, Y., Mo, D., Zhang, L., Boswell, M., & Rozelle, S. (2016). The impact of integrating ICT with teaching: Evidence from a randomized controlled trial in rural schools in China. Computers & Education, 96, 1–14. https://doi.org/10.1016/j.compedu.2016.02.005

Bando, R., Gallego, F., Gertler, P., Romero Fonseca, D. (2017). Books or laptops? The effect of shifting from printed to digital delivery of educational content on learning. Economics of Education Review. 61, 162-173. 10.1016/j.econedurev.2017.07.005

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. The Quarterly Journal of Economics, 122(3), 1235–1264. https://doi.org/10.1162/qjec.122.3.1235

Barrera-Osorio, F., & Linden, L. (2009). The use and misuse of computers in education: Evidence from a randomized experiment in Colombia. Policy Research Working Paper, WPS4836. http://documents.worldbank.org/curated/en/346301468022433230/The-use-and-misuse-of-computers-in-education-evidence-from-a-randomized-experiment-in-Colombia

Beg, S., Lucas, A., Halim, W., & Saif, U. (2019). Beyond the Basics: Improving Post-Primary Content Delivery through Classroom Technology (No. w25704; p. w25704). National Bureau of Economic Research. https://doi.org/10.3386/w25704

Berkhout, E., Pradhan, M., Rahmawati, Suryadarma, D., Swarnata, A. 2020. From Cheating to Learning: An Evaluation of Fraud Prevention on National Exams in Indonesia. RISE Working Paper Series. 20/046. https://doi.org/10.35489/BSG- RISEWP_2020/046.

Berlinski, S., & Busso, M. (2017). Challenges in educational reform: An experiment on active learning in mathematics. Economics Letters, 156, 172–175. https://doi.org/10.1016/j.econlet.2017.05.007

Berlinski, S., Busso, M., Dinkelman, T., & Martinez, C. (2016). Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile. Working Paper. https://www.povertyactionlab.org/sites/default/files/publications/726_%20Reducing-Parent-School-information-gap_BBDM-Dec2016.pdf

Bet, G., Cristia, J., & Ibarrarán, P. (2014). The effects of shared school technology access on students' digital skills in Peru. IDB Working Paper Series No. IDB-WP--476. https://publications.iadb.org/publications/english/document/The-Effects-of-Shared-School-Technology-Access-on-Students-Digital-Skills-in-Peru.pdf

Bettinger, E., Kardanova, E., Fairlie, R., Loyalka, P., Kapuza, A., & Zakharov, A. (2020). "Does EdTech substitute for traditional learning? Experimental estimates of the educational production function. Stanford Institute for Economic Policy Research (SIEPR). Working paper No. 20-010.

Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O., & Cruz-Aguayo, Y. (2015). One Laptop per Child at Home: Short-Term Impacts from a Randomized Experiment in Peru. American Economic Journal: Applied Economics, 7(2), 53–80. https://doi.org/10.1257/app.20130267

Bianchi, N., Lu, Y., & Song, H. (2019). The Effect of Computer-Assisted Learning on Students' Long-Term Development. https://ssrn.com/abstract=3309169 or http://dx.doi.org/10.2139/ssrn.3309169

Blimpo, M. P., Gajigo, O., Owusu, S., Tomita, R., & Xu, Y. (2020). Technology in the classrrom and learning in secondary schools. Unpublished manuscript.

Böhmer, B., Burns, J., & Crowley, L. (2014). Testing Numeric: Evidence from a randomized controlled trial of a computer-based mathematics intervention in Cape Town High Schools. University of Cape Town.

Borzekowski, D. L. G. (2018). A quasi-experiment examining the impact of educational cartoons on Tanzanian children. Journal of Applied Developmental Psychology, 54, 53–59. https://doi.org/10.1016/j.appdev.2017.11.007

Borzekowski, D. L. G., & Henry, H. K. (2011). The impact of Jalan Sesama on the educational and healthy development of Indonesian preschool children: An experimental study. International Journal of Behavioral Development, 35(2), 169–179. https://doi.org/10.1177/0165025410380983

Borzekowski, D. L. G., Lando, A. L., Olsen, S. H., & Giffen, L. (2019a). The Impact of an Educational Media Intervention to Support Children's Early Learning in Rwanda. International Journal of Early Childhood, 51(1), 109–126. https://doi.org/10.1007/s13158-019-00237-4

Borzekowski, D., Singpurwalla, D., Mehrotra, D. & Howard, D. (2019b). The impact of Galli Galli Sim Sim on Indian preschoolers. Journal of Applied Developmental Psychology, 64. Article number 101054. https://doi.org/10.1016/j.appdev.2019.101054

Brown, F., Farag, A., Hussein, F., Miller, L., Radford, K., Abbadi, A.A., Neijenhuijs, K., Oprins, E., De Hoop, T., Turner, J., Jetten, A., & Jordans, M. (Forthcoming). Can't Wait to Learn: A quasi-experimental evaluation of a digital game-based learning programme for out of school children in Sudan. Journal of Development Effectiveness.

Büchel, K., Jakob, M., Kühnhanss, C., Steffen, D., Brunetti, A. (2020). The relative effectiveness of teachers and learning software: evidence from a field experiment in El Salvador. University of Bern Social Sciences Working Paper No. 36

Cardim, J., Molina Millán, T., and Vicente, P. (2019). Technology-Aided Instruction in Primary Education: Experimental Evidence from Angola [Working Paper]. https://custom.cvent.com/4E741122FD8B4A1B97E483EC8BB51CC4/files/Event/ca4fa1845a194f7fa5fed5711b3c0a8f/ce84c22d75cd42748fd5feca3991f2a8.pdf

Carrillo, P. E., Onofa, M., & Ponce, J. (2011). Information Technology and Student Achievement: Evidence from a Randomized Experiment in Ecuador. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1818756

Chong, A., Gonzalez-Navarro, M., Karlan, D., Valdivia, M. (2020). The World Bank Economic Review. 34(2), 371-392. doi: 10.1093/wber/lhy031

Cilliers, J. Fleisch, B. Kotze, J., Mohohlwane, N., Thulare, T., & Taylor, S. (2020). Can Virtual Replace In-person Coaching? Experimental Evidence on Teacher Professional Development and Student Learning in South Africa. Research on Improving Systems of Education (RISE). https://www.riseprogramme.org/sites/www.riseprogramme.org/files/inline-files/Virtual%20vs%20InPerson%20Coaching%20Working%20Paper.pdf

Cristia, J., Czerwonko, A., & Garofalo, P. (2010). Does ICT increate years of education? Evidence from Peru. Inter-American Development Bank Working Paper: OVE/WP-01/10.

Cristia, J., Ibarrarán, P., Cueto, S., Santiago, A., & Severín, E. (2017). Technology and Child Development: Evidence from the One Laptop per Child Program. American Economic Journal: Applied Economics, 9(3), 295–320. https://doi.org/10.1257/app.20150385

De Hoop, T., Brudevold-Newman, A., Ring, H., Riyadh, H., Turner, J., Hunt, K., Rothbard, V., Radford, K., Twissi, Z., Jordans, M.J.D., & Brown, F. (2020). Can comprehensive gaming-based technology in formal education programs improve learning and psychosocial outcomes? Evidence from a quasi-experimental study in Jordan (Work in Progress).

de Hoop, T., Ring, H., Siwach, G., Dias, P., Tembo, G., Rothbard, V., and Toungui, A. (2020). Impact of e-Learning Technology and Activity-Based Learning on Learning Outcomes: Experimental Evidence from Community Schools in Rural Zambia.

de Melo, G., Machado, A., & Miranda, A. (2014). The Impact of a One Laptop per Child Program on Learning: Evidence from Uruguay. IZA, Discussion Paper No. 8489.

Derksen, L., Leclerc, C., and Souza, P. C. (2020). Searching for Answers: The Impact of Student Access to Wikipedia [Working Paper]. https://warwick.ac.uk/fac/soc/economics/research/centres/cage/manage/publication s/450-2019_souza.pdf Deschênes, S., and Hotte, R.

Duflo, E., Hanna, R., & Ryan, S. P. (2012). Incentives Work: Getting Teachers to Come to School. American Economic Review, 102(4), 1241–1278. https://doi.org/10.1257/aer.102.4.1241

Dustan, A., Maldonado, S., & Hernandez-Agramonte, J.M. (2019). Motivating bureaucrats with non-monetary incentives when state capacity is weak: Evidence from large-scale field experiments in Peru. Working paper.

Freeman, B., & Hawkins, R. (2017). Evoke. Developing skills in youth to solve the world's most complex problems: randomized impact evaluation findings. World Bank Education, Technology & Innovation: SABER-ICT Technical Paper Series No. 19.

Gaduh, A., Pradhan, M., Priebe, J., & Susanti, D. (2020). Scores, Camera, Action? Incentivizing Teachers in Remote Areas. Research on Improving Systems of Education (RISE). https://doi.org/10.35489/BSG-RISE-WP_2020/035

Gambari, A. I., Shittu, A. T., Daramola, F. O., & James, M. (2016). Effects of video-based cooperative, competitive and individualized instructional strategies on the performance of senior secondary schools students in geometry. Malaysian Online Journal of Educational Sciences. 4(4).

Habyarimana, J., & Jack, W. (2018). High Hopes: experimental evidence on saving and the transition to high school in Kenya. Georgetown University Initiative on Innovation, Development and Evaluation Working Paper Series No. 4.

Habyarimana, J., & Sabarwal, S. (2018). Re-Kindling Learning: EReaders in Lagos. Policy Research Working Paper;No. 8665. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/30987

He, F., Linden, L., & MacLeod, M. (2008). How to Teach English in India: Testing the Relative Productivity of Instruction Methods within the Abstract: Pratham English Language Education Program. Working Paper. https://pdfs.semanticscholar.org/54de/5782f2c61a6fe8cb261c1e5e05904694d6b1.pdf

Hirshleifer, S.R. (2016). Incentives for effort of outputs? A field experiment to improve student performance. Working Papers 201701, University of California at Riverside, Department of Economics.

Ito, H., Kasai, K., & Nakamuro, M. (2019). Does Computer-Aided Instruction Improve Children's Cognitive and Non-Cognitive Skills? Evidence from Cambodia. Discussion Papers 19040, Research Institute of Economy, Trade and Industry (RIETI). https://ideas.repec.org/p/eti/dpaper/19040.html

Jere-Folotiya, J., Chansa-Kabali, T., Munachaka, J.C., Sampa, F., Yalukanda, C., Westerholm, J., Serpell, R., Lyytinen, H. (2014). The effect of using a mobile literacy game to improve literacy levels of grade one students in Zambian schools. Education Technology Research and Development. 62, 417–436. https://doi.org/10.1007/s11423-014-9342-9

John, I. E., Udofia, N.-A., Udoh, N. A., and Anagbogu, M. A. (2016). Development of ecareer guidance programme for secondary schools in Akwa Ibom State. Educational

Media International, 53(4), 298–312. https://doi.org/10.1080/09523987.2016.1254886"

Johnston, J., & Ksoll, C. (2017). Effectiveness of Interactive Satellite-Transmitted Instruction: Experimental Evidence from Ghanaian Primary Schools. Working Paper 17–08. CEPA Working Papers. Stanford, CA: Stanford Center for Education Policy Analysis. https://eric.ed.gov/?id=ED579066

Kho, K., Lakdawala, L., & Nakasone, E. (2018). Impact of Internet Access on Student Learning in Peruvian Schools. https://mwiedc.aae.wisc.edu/schedule/papers/283/

Kotze, J., Fleisch, B., & Taylor, S. (2019). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. International Journal of Educational Development, 66, 203–213. https://doi.org/10.1016/j.ijedudev.2018.09.004

Lai, F., Luo, R., Zhang, L., Huang, X., & Rozelle, S. (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. https://doi.org/10.1016/j.econedurev.2015.03.005

Lai, F., Zhang, L., Bai, Y., Liu, C., Shi, Y., Chang, F., & Rozelle, S. (2016). More is not always better: Evidence from a randomised experiment of computer-assisted learning in rural minority schools in Qinghai. Journal of Development Effectiveness, 8(4), 449–472. https://doi.org/10.1080/19439342.2016.1220412

Lai, F., Zhang, L., Hu, X., Qu, Q., Shi, Y., Qiao, Y., Boswell, M., & Rozelle, S. (2013). Computer assisted learning as extracurricular tutor? Evidence from a randomised experiment in rural boarding schools in Shaanxi. Journal of Development Effectiveness, 5(2), 208–231. https://doi.org/10.1080/19439342.2013.780089

Lehrer, K., Mawoyo, M., and Mbaye, S. (2019). The Impacts of Interactive Smartboards on Learning Achievement in Senegalese Primary Schools (3ie Grantee Final Report). https://developmentevidence.3ieimpact.org/search-result-details/impact-evaluationrepository/the-impacts-of-interactive-smartboards-on-learning-achievement-insenegalese-primary-schools/7721

Linden, L. (2008). Complement or substitute? The effect of technology on student achievement in India. Working Paper (Numbered Series), 44863. http://documents.worldbank.org/curated/en/804371468034237060/Complement-or-substitute-The-effect-of-technology-on-student-achievement-in-India

Lysenko, L., Abrami, P. C., Wade, C. A., Marsh, J. P., WaGioko, M., & Kiforo, E. (2019). Promoting young Kenyans' growth in literacy with educational technology: A tale of two years of implementation. International Journal of Educational Research. doi:10.1016/j.ijer.2019.02.013

Ma, Y., Fairlie, R., Loyalka, P., & Rozelle, S. (2020). Isolating the "Tech" from EdTech: Experimental Evidence on Computer Assisted Learning in China (No. w26953; p. w26953). National Bureau of Economic Research. https://doi.org/10.3386/w26953

Malamud, O., & Pop-Eleches, C. (2011). Home computer use and the development of human capital. The Quarterly Journal of Economics, 126, 987–1027.

Malamud, O., Cueto, S., Cristia, J., & Beuermann, D. W. (2019). Do children benefit from internet access? Experimental evidence from Peru. Journal of Development Economics, 138, 41–56. https://doi.org/10.1016/j.jdeveco.2018.11.005

Mensch, B., & Haberland, N. (2018). GirlsRead! Zambia DREAMS Innovation. ISRCTN. https://www.popcouncil.org/uploads/pdfs/2018PGY_GirlsRead_Brief.pdf

Meza-Cordero, J. A. (2017). Learn to Play and Play to Learn: Evaluation of the One Laptop per Child Program in Costa Rica: Learn to Play and Play to Learn. Journal of International Development, 29(1), 3–31. https://doi.org/10.1002/jid.3267

Mo, D., Swinnen, J., Zhang, L., Yi, H., Qu, Q., Boswell, M., & Rozelle, S. (2013). Can One-to-One Computing Narrow the Digital Divide and the Educational Gap in China? The Case of Beijing Migrant Schools. World Development, 46, 14–29. https://doi.org/10.1016/j.worlddev.2012.12.019

Mo, D., Zhang, L., Luo, R., Qu, Q., Huang, W., Wang, J., Qiao, Y., Boswell, M., & Rozelle, S. (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. https://doi.org/10.1080/19439342.2014.911770

Mo, D., Zhang, L., Wang, J., Huang, W., Shi, Y., Boswell, M., & Rozelle, S. (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).

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. American Economic Review, 109(4), 1426–1460. https://doi.org/10.1257/aer.20171112

Naik, G., Chitre, C., Bhalla, M., & Rajan, J. (2020). Impact of use of technology on student learning outcomes: Evidence from a large-scale experiment in India. World Development, 127, 104736. https://doi.org/10.1016/j.worlddev.2019.104736

Näslund-Hadley, E., Parker, S., & Hernandez-Agramonte, J. M. (2014). Fostering early math comrehension: Experimental evidence from Paraguay. Global Education Review., 1(4):135-54. https://eric.ed.gov/?id=EJ1055163

Navarro-Sola, L. (2019). Secondary School Expansion through Televised Lessons: The Labor Market Returns of the Mexican Telesecundaria. Working Paper. https://laianaso.github.io/laianavarrosola.com/Navarro-Sola_JMP.pdf

Neilson, C., Berry, J, Coffman, L., Morales, D. (2014). Learning the Value of Education in the Dominican Republic. Policy report: Milestone 13. The Abdul Latif Jameel Poverty Action Lab, Latin America & the Caribbean.

Neilson, C., Gallegos, F., & Molina, O. (2018). Decidiendo para un futuro mejor: Informando sobre los retornos monetarios y sociales de la educación. Policy report: Evidencias MineduLAB No. 06.

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 od Educational Development. 49. 204-2014. http://dx.doi.org/10.1016/j.ijedudev.2016.03.006

Pitchford, N. J. (2015). Development of early mathematical skills with a tablet intervention: A randomized control trial in Malawi. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.00485

Pitchford, N. J., Chigeda, A., Hubber, P. J. (2019). Interactive apps prevent gender discrepancies in early-grade mathematics in a low-income country in sub-Saharan Africa. Developmental Science. Special Issue Article. DOI: 10.1111/desc.12864

Riley, E. (2017). Role models in movies: the impact of Queen of Katwe on students' educational attainment. CSAE Working Paper Series 2017-13, Centre for the Study of African Economies, University of Oxford.

Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., Grau, V., Lagos, F., López, X., López, V., Rodriguez, P., & Salinas, M. (2003). Beyond Nintendo: Design and assessment of educational video games for first and second grade students. Computers & Education, 40(1), 71–94. https://doi.org/10.1016/S0360-1315(02)00099-4

Seo, H. K. (2017). Do School Electrification and Provision of Digital Media Deliver Educational Benefits? First-Year Evidence from 164 Tanzanian Secondary Schools. Working Paper E-40308-TZA-2. London: International Growth Centre. https://www.theigc.org/wp-content/uploads/2017/10/IGC_Final-report_cover2.pdf

Vakis, R., & Farfan, G. (2018). Envío de mensajes de texto para incrementar la motivación y satisfacción docente. Policy report: Evidencias MineduLAB No. 04.

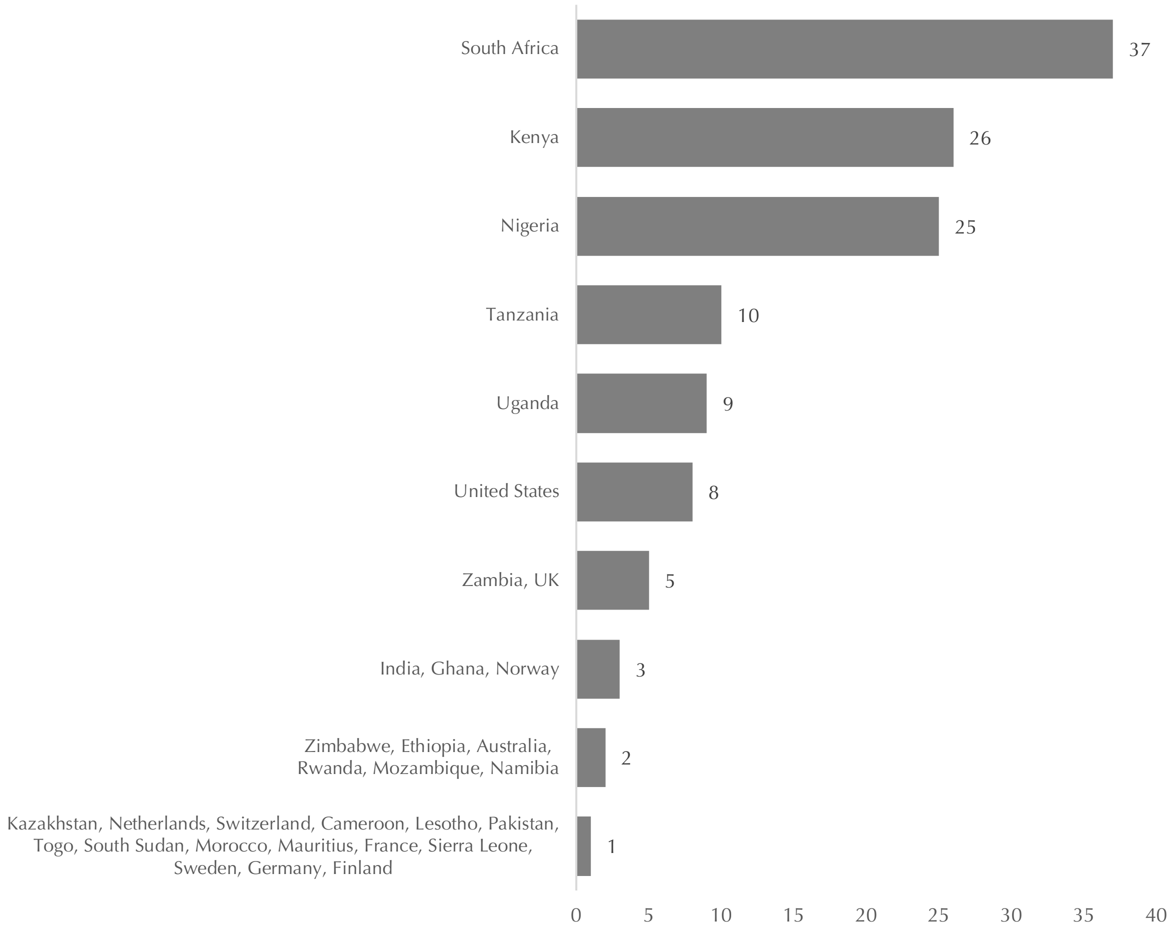
Wennersten, M., Quraishy, Z. B., & Velamuri, M. (2015). Improving student learning via mobile phone video content: Evidence from the BridgeIT India project. International Review of Education, 61(4), 503–528. https://doi.org/10.1007/s11159-015-9504-y

Wolf, S., Aber, J. L., Behrman, J. R., Tsinigo, E. (2018). Experimental Impacts of the “Quality Preschool for Ghana” Interventions on Teacher Professional Well-being, Classroom Quality, and Children’s School Readiness. Journal of Research on Educational Effectiveness. 0(0). 1-28. https://doi.org/10.1080/19345747.2018.1517199

Yanguas, M. L.. (2020). Technology and educational choices: evidence from a one-laptop-per-child program. Economics of Education Review. 76. Article number 101984. https://doi.org/10.1016/j.econedurev.2020.101984

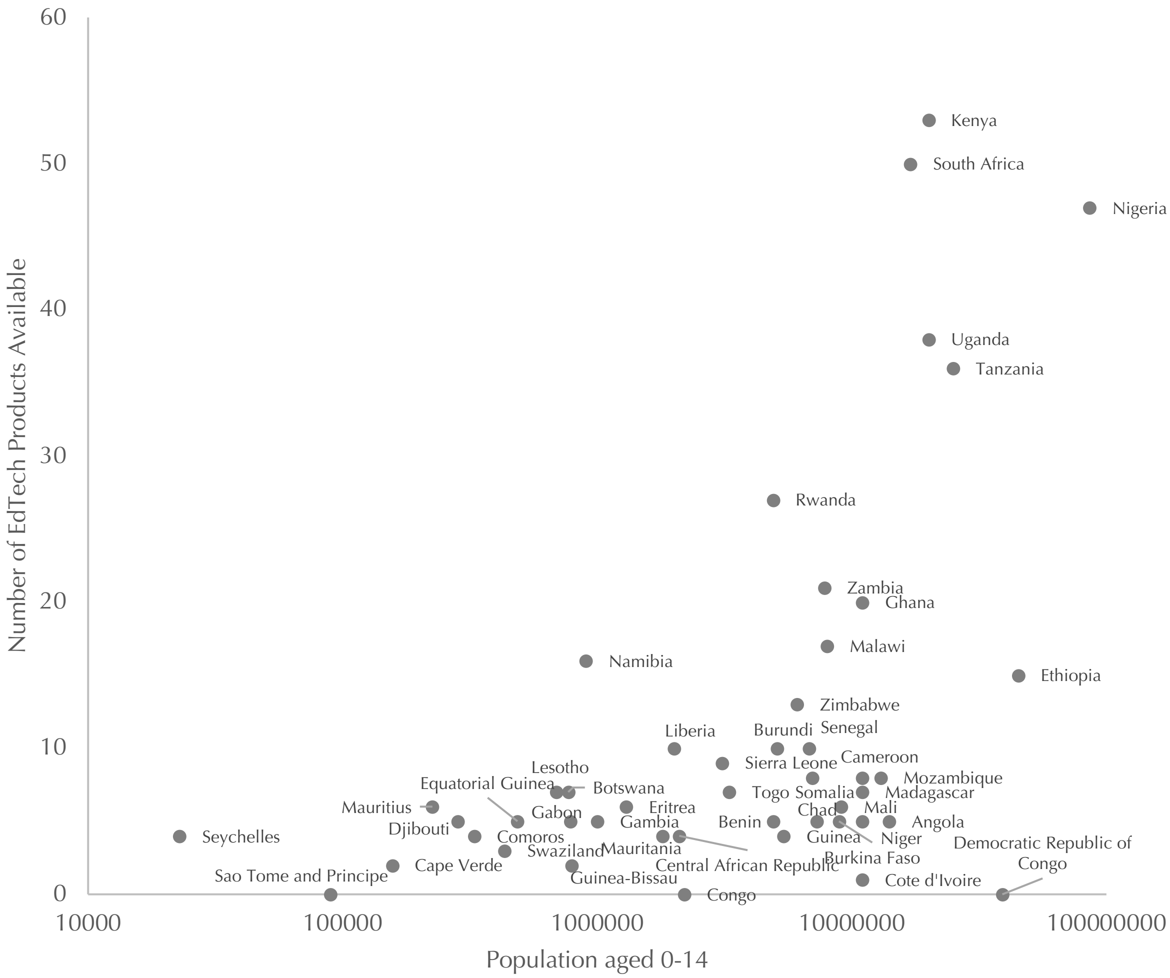
Appendix A: Additional figures

Figure A1: number of EdTech firms by country of origin



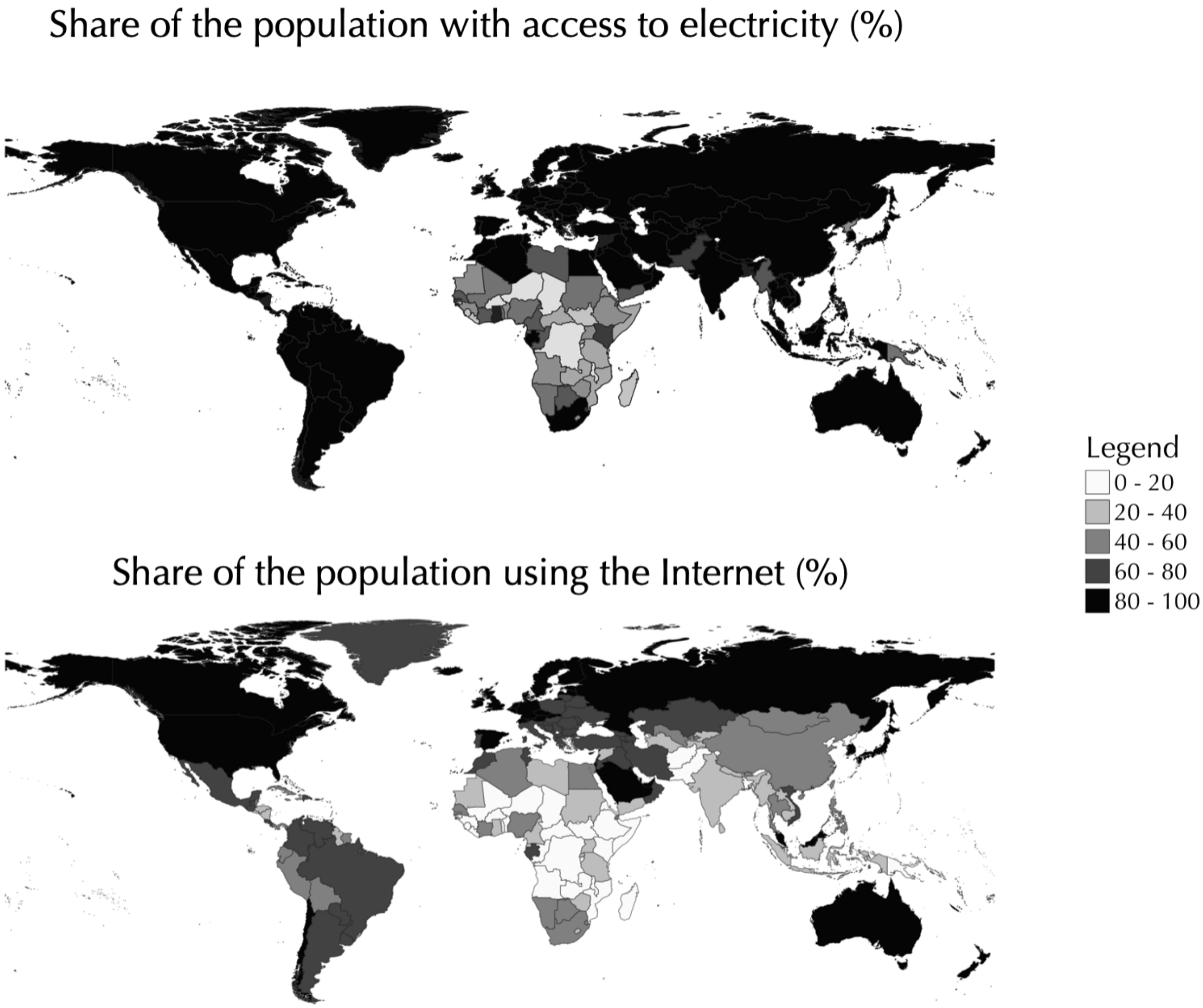
Notes: the data, code, and ideation for this graph were kindly shared by Lee Crawfurd, all of which were first used in his May 2020 blog post (Crawfurd, 2020).

Figure A2: number of EdTech firms by country of origin compared to population aged 0-14, for Sub-Saharan African countries



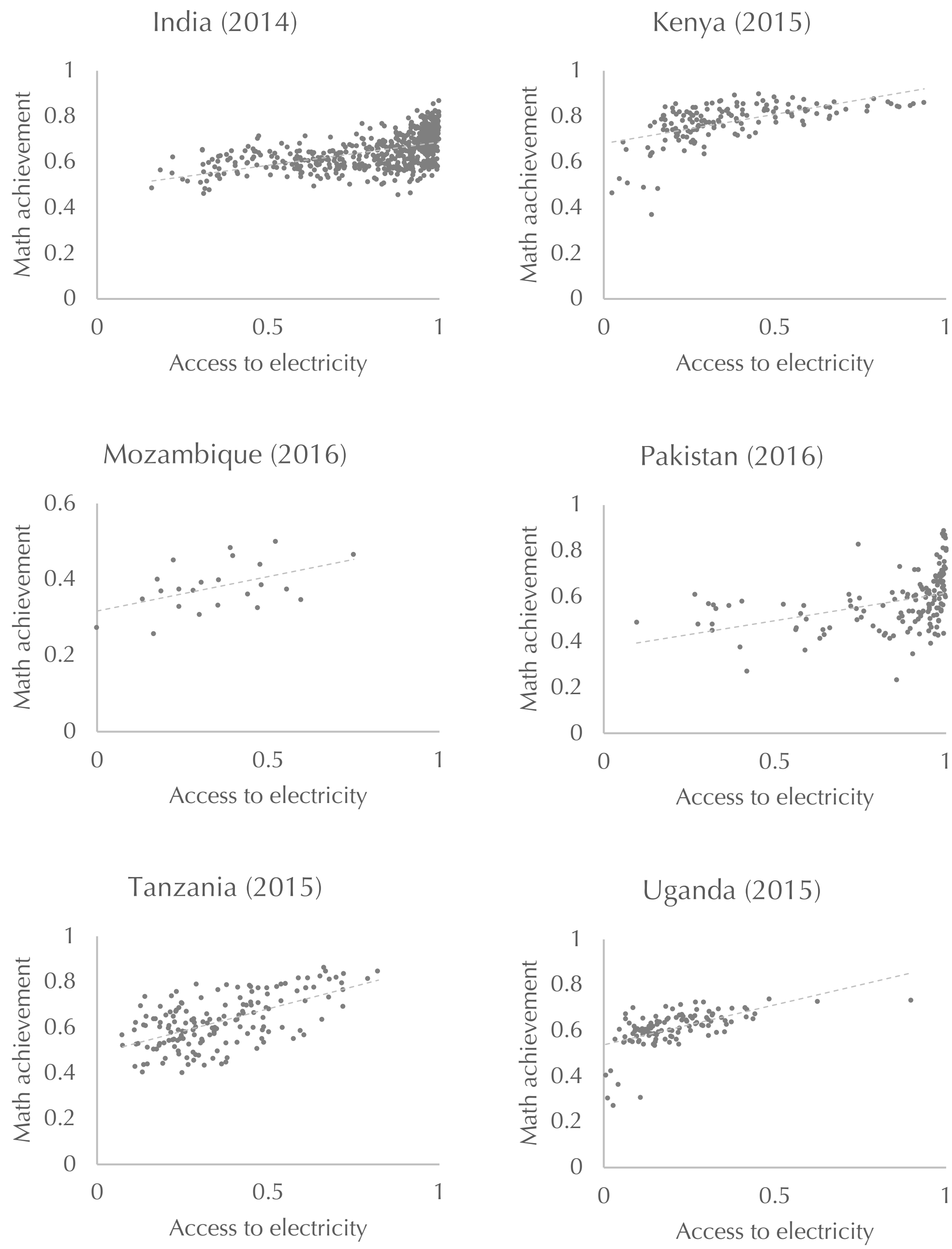
Notes: the data, code, and ideation for this graph were kindly shared by Lee Crawfurd, all of which were first used in his May 2020 blog post (Crawfurd, 2020).

Figure A3: access to electricity and internet around the world



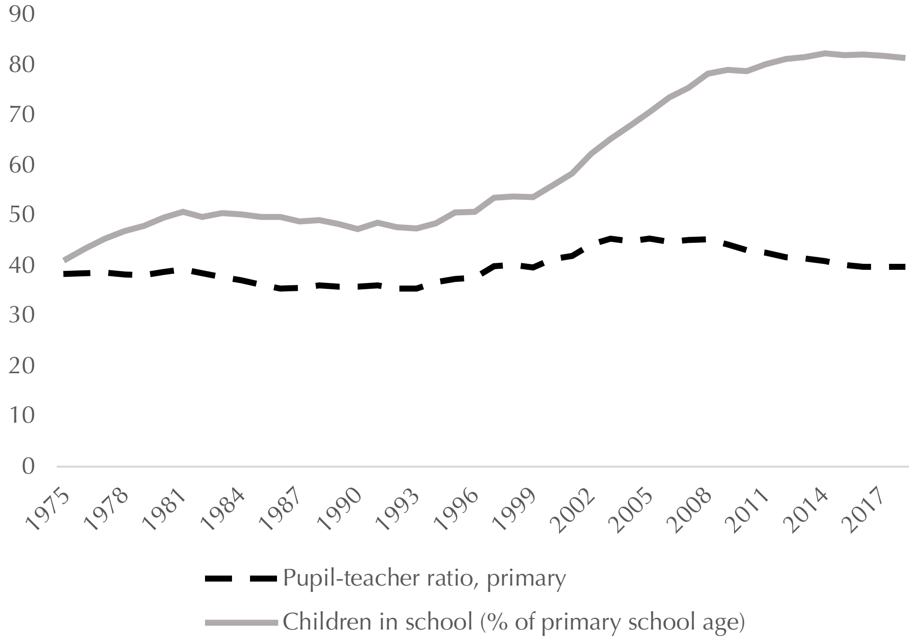
Notes: the data is from the World Development Indicators ("Access to electricity (% of population)" and "Individuals using the Internet (% of population)"). Each country displays the latest value available in the raw data.

Figure A4: relationship between household electrification and math achievement at the district-level



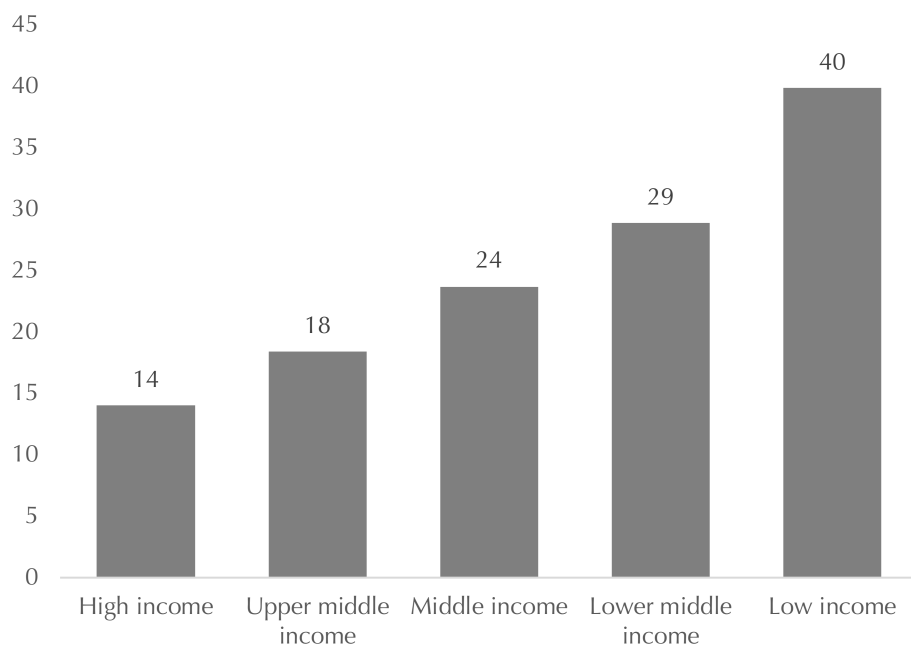
Notes: the horizontal axis displays the district-level average of household with access to electricity, and on the vertical axis the mean proficiency in an internationally-comparable basic math exam. Data for Uganda, Kenya, and Tanzania comes from the Uwezo nationwide household survey from 2015 (Twaweza, 2015). The data for Pakistan and India comes from the nationwide ASER surveys from 2016 and 2014 respectively (ASER, 2016 and ASER, 2014). The data for Mozambique comes from the 2016 pilot of "Todos Pelas Crianças" (TPC) in the Nampula province (TPC, 2016). Aggregation at the district-level, and harmonization across countries were the author's own elaboration.

Figure A5: comparison of primary school enrollment in low income countries, and primary school pupil-teacher ratio in low income countries



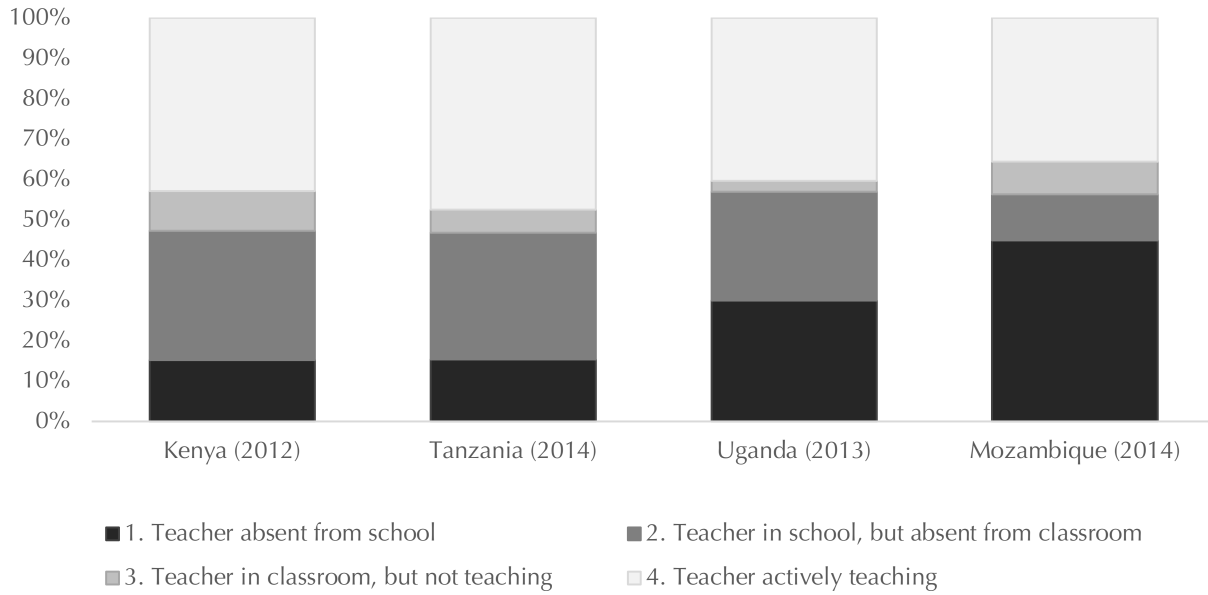
Notes: numbers from World Development Indicators, using the "Pupil-teacher ratio, primary" indicator and the inverse of "Children out of school (% of primary school age)"

Figure A6: primary school pupil-teacher ratio by income classification



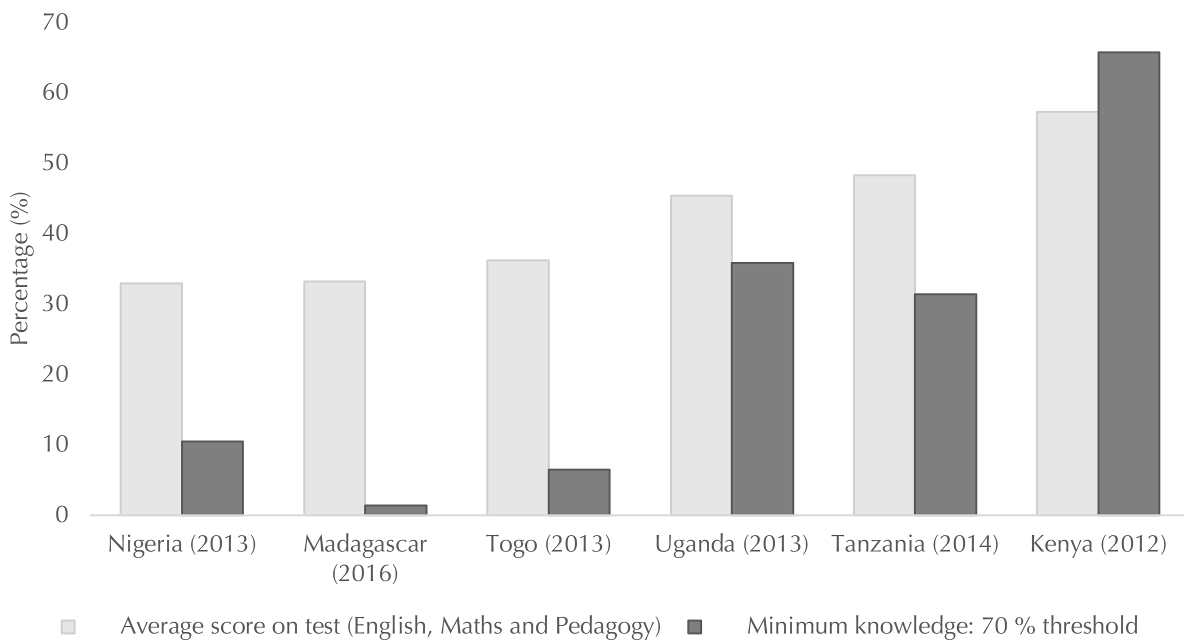
Notes: numbers from World Development Indicators, using the "Pupil-teacher ratio, primary" indicator

Figure A7: teacher time allocation for selected Eastern African countries



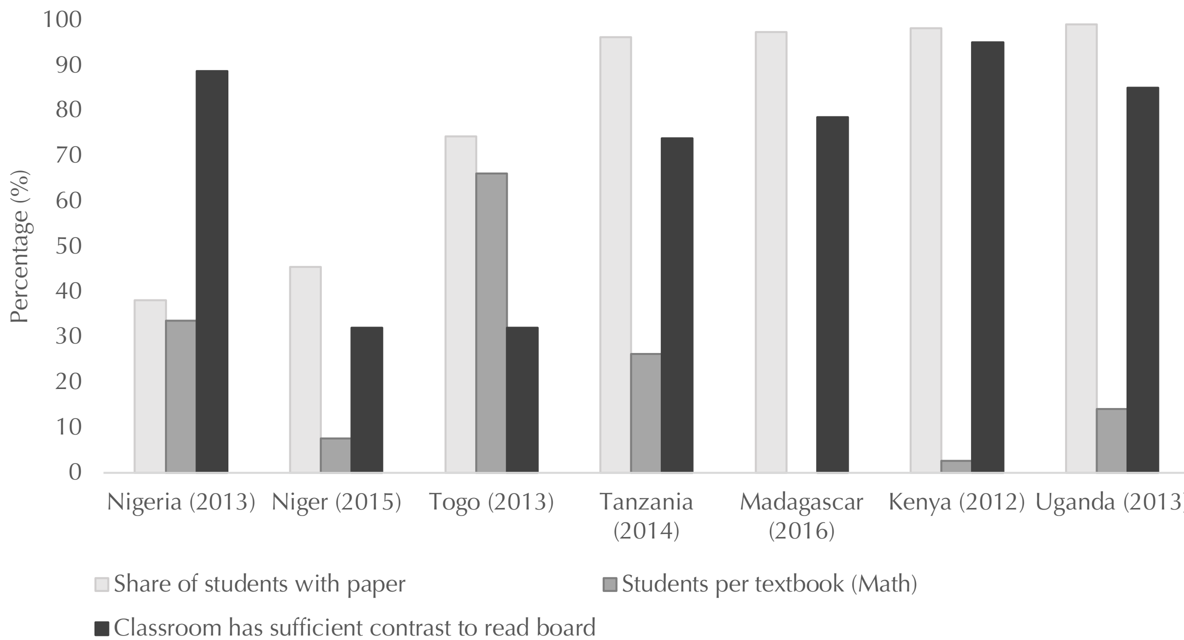
Notes: numbers from the Service Delivery Indicators (SDI) by the World Bank. Calculations using the “Absence from school”, “Absence from classroom”, and “Time spent teaching (minutes)” indicators.

Figure A8: measures of teacher knowledge for selected African countries



Notes: numbers from the Service Delivery Indicators (SDI) by the World Bank

Figure A9: measures of classroom input availability for selected African countries



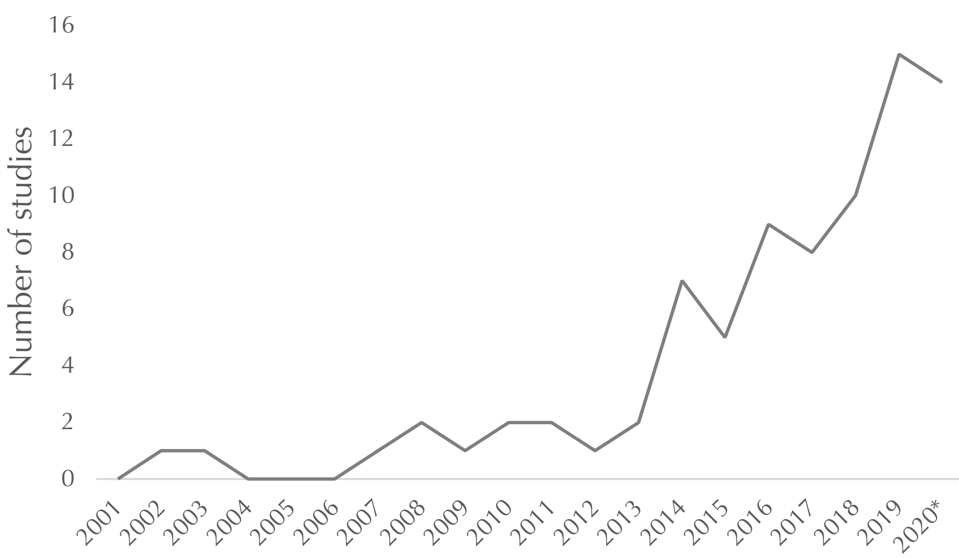
Notes: numbers from the Service Delivery Indicators (SDI) by the World Bank

Appendix B: Further methodological considerations

1. Other study features which were not part of the inclusion and exclusion criteria

In the spirit of following Escueta et al. (2020) to be as inclusive of high-quality, relevant studies as possible, this review makes the explicit decision to not filter papers by any other criterion not mentioned above. Among the potential filtering criteria that did not play a role in the selection of the core studies, the time of publication is one of the most salient ones: there was no minimum year for the inclusion of a paper in the review, especially given that the oldest study found dates back to only 2002. Since not all studies have been published in an academic journal, the date for each study refers to either the date of publication in a peer-reviewed journal, or the date on the latest draft found for each study. Figure A10 below provides a sense of the temporal distribution of studies: interestingly, the number has increased significantly since 2013, reaching 15 studies only for 2019. This time trend highlights the growing interest in the field of education in developing settings from researchers, and the further need for a compilation of all existing evidence to date.

Figure A10: number of studies in this review by the year of their latest draft



* The 2020 value is as of September, 2020

Another feature which was not used to filter studies was the targeted outcomes and stakeholders. While 89% of all core studies either only targeted learning outcomes, or had it as one of its main outcomes of interests, there were other important outcomes studied, such as school enrollment, dropout rates, sexual health behaviors, and motivation. Similarly, the review was open to studies targeting all kinds of educational stakeholders. A vast majority of the interventions (83%) were student-facing and targeted students in grades 1-12, but there were other groups studied included such pre-K students, university students, teachers, civil servants and parents.

Finally, the scale of the technology used did not play a role in the selection of the studies. The studied technology could be a large national rollout requiring large investments such as telesecundarias in Mexico (Navarro-Sola, 2019), or lower-touch text message interventions in Peru such as in Neilson et al. (2018a, 2018b). Similarly, there was no restriction on the sample size for the study, ranging from a few hundred observations like in Pitchford (2015), Mo et al. (2013), or Böhmer et al. (2014), to upwards of 100,000 in an experimental set up such as Neilson et al. (2018b) and almost 900,000 in a quasi-experimental setup (Navarro-Sola, 2019).

2. Search methods

The search for papers that make up the set of core studies was at the forefront of the evidence-gathering process for this review. The first round of searches was within repositories of peer-reviewed journals and databases such as EconLit, EconPapers, and Google Scholar, where multiple combinations of words related to the scope of this review[36] were searched. Furthermore, I looked for the same terms in the AEA Trial Registry for any trials that may have finished already. Next, I looked in the working paper repositories of well-known organizations that routinely produce education-related research as the World Bank, the Interamerican Development Bank, the EdTech Hub, NBER, the RISE Programme, Annenberg Institute, J-PAL, and IPA. I also used back- and forward tracing of citations from four highly cited and/or comprehensive papers: Muralidharan et al. (2019), Sampson et al. (2019), Escueta et al. (2020), Evans and Mendez Acosta (2020), and World Bank (2018). After identifying an initial set of papers through these methods, I forward-traced papers through the literature review sections of these papers, and the papers that they cite. I then backward-tracked, i.e. searched other papers that cited these studies, each of these papers through the Google Scholar feature for this process ("Cited by"). After completing this process, I iterated through the process of back- and forward-tracing papers until no additional papers were located. While there is no guarantee that all studies that meet the four main criteria are included in

[36] The actual terms searched were "EdTech", "ed-tech", "Ed Tech", "Technology education", "Technology in education", "ICT in education", "SMS education", "Computers education", "Laptops education", "Technology instruction", "Technology school" all by themselves, and then combining them with "developing countries", "Latin America", "Africa", "Sub-Saharan Africa", and "India".

the set of core studies, great lengths were covered to ensure that the review was as extensive as possible.

Appendix C: Non-comprehensive list of upcoming EdTech studies

Table A1: studies with a considerable EdTech component but for which there is no write up publicly available by the time this review was completed

Researchers	Context	Project Title	Intervention	Source
Guilherme Lichand and Sharon Wolf	Côte d'Ivoire	Evaluating the Impact of Text and Audio Messages for Parents and Teachers in Côte d'Ivoire	Text and audio messages for parents either with or without messages to teachers to increase attendance in school.	https://www.poverty-action.org/study/evaluating-impact-text-and-audio-messages-parents-and-teachers-côte-d'ivoire
Emma Näslund-Hadley and Juan Manuel Hernandez Agramonte	Paraguay	The Effects of Interactive Radio Instruction for Science Education in Paraguay	Interactive audio instruction ("IRI") curriculum for early childhood education, particularly in science. Following success of similar project in Math.	https://www.poverty-action.org/study/effects-interactive-radio-instruction-science-education-paraguay
Juan Manuel Hernandez Agramonte and Mercedes Mateo-Berganza	Uruguay	The Impact of Text-Message Nudges on Preschool Attendance in Uruguay	Behaviorally-informed SMS messages to parents informing them of the importance of early childhood education to encourage preschool attendance.	https://www.poverty-action.org/study/impact-text-message-nudges-preschool-attendance-uruguay
Emma Näslund-Hadley, Juan Manuel Hernandez Agramonte, and Elena Arias Ortiz	Costa Rica	Using a Robot to Improve Young Children's Math and Programming Skills in Costa Rica	The Pensalo program introduces an intelligent robot named "Albert" that 4 and 5 year old students program by scanning a series of flash cards with instructions that use mathematical and numerical concepts.	https://www.poverty-action.org/study/using-robot-improve-young-children's-math-and-programming-skills-costa-rica
Emma Näslund-Hadley and Juan Manuel Hernandez Agramonte	Colombia	The Effects of a Multimedia Preschool STEM Education Program in Colombia	The program includes a web series, computer games, and interactive posters that teach children STEM-related conceptsm, and is facilitated by "community mothers though teaching guide, video tutorials, and a structured lesson planon 4-5 year olds.	https://www.poverty-action.org/study/effects-multimedia-preschool-stem-education-program-colombia
Bruno Ferman, Lycia Lima, Flávio Riva	Brazil	The Impact of Automated Writing Evaluation on Learning and Access to College in Brazil	Evaluation of whether programs using natural language processing, and machine-learning algorithm to score and comment on easays can improve learning and increase access to college for secondary students in public schools in Brazil.	https://www.povertyactionlab.org/evaluation/impact-automated-writing-evaluation-learning-and-access-college-brazil
Bruno Crépon, Igor Asanov, Diego d'Andria, Thomas Astebro, Guido Buenstorf, Francisco Flores, Mona Mensmann, Mathis Schulte, David McKenzie	Ecuador	The impact of an online entrepreneurial mindset training for youth in Ecuador	Online-based psychology-based entrepreneurial mindset training paired with either negotiations skills or scientific skills training, and mentoring.	https://www.povertyactionlab.org/evaluation/impact-online-entrepreneurial-mindset-training-youth-ecuador
Adrienne Lucas, Sabrin Beg, and	Pakistan	Screen Time: Tablets with Interactive	Using an RCT among grade 6 students in Punjab, Pakistan, we tested the effect of providing tablets	Information from the authors

Samantha Sweeney		Textbooks Did Not Increase Learning	with interactive textbooks to students on their achievement in math and science. We found no evidence that the intervention improved test scores 3 months after implementation.	
Alejandro Ganimian, Karthik Muralidharan, and Andy de Barros	India	Do Students Benefit from Personalized Learning? Experimental Evidence from India	Personalized instruction delivered by computer-assisted learning software. Comparison of software that provides only grade-appropriate activities, with fully and partially customized version of program, as well as a remedial version of it.	https://www.socialscienceregistry.org/trials/2459/history/21859

Online Tables

Online Table 1: summary of studies included within the "Access to technology" category

Study	Intervention	Context	Target grade and outcomes	Sample	Findings	Cost
Angrist and Lavy (2002)	Program "Tomorrow-98". Target student-computer ratio of 10:1 in all schools. Additional teacher training to integrate computers to instruction. Program assignment at the school-level.	Israel	Grades 4 and 8. 122 schools, targeted at elementary and middle schools throughout Israel.	4,779 4th graders, 3,196 8th graders	Grade 4: -0.4 to -0.3 SD in Math, no effects in Hebrew. No effects in grade 8 across most models.	USD 3000 per computer, with 40 computers per school.
Bando et al. (2017)	Replacement of traditional textbooks with laptops. Randomization at school-level.	Honduras	Grades 3 and 6. 271 elementary schools throughout the country.	9,600	No effects.	Net cost of USD 48 per student, per year.
Barrera-Osorio and Linden (2009)	Program "Computadores para Educar". 15 computers per school to support children's language. 20-month long training for teachers. Randomization at school-level.	Colombia	Grades 3-9, 97 public schools with 80 or more students. Six school districts.	5,201	No effects.	Not specified.
Bet et al. (2014)	Propensity score matching groups with similar observable educational inputs but different intensity in computer access.	Peru	Grade 9, 202 schools.	4,897	No effects in math or language, 0.3 SD in digital skills.	Not specified.
Beuermann et al. (2015)	Program "One Laptop per Child". Four laptops, one per student, randomly distributed in each class for use at home. Each computer included applications such as educational games, programming environments, and an encyclopedia. Seven weekly training sessions. Randomization at the student-level within classes in treatment schools.	Peru	Grade 2, 28 schools, Public schools in Lima.	2,734	No effects on achievement level. Increased computer proficiency in treated students.	USD 188 per laptop.
Cardim et al. (2019)	Evaluation of "ProFuturo" intervention. The program includes the distribution of suitcases with tablets, a computer for the teacher and a projector. Randomization at the school-level.	Angola	Grades 4-6. 42 Catholic schools in Luanda.	2,460	No effects in learning, increased familiarity with technology.	Not specified.
Cristia et al. (2010)	ICT regional package including the lay-out of the electrical infrastructure, 10 computers and the installation of a network. These schools entered the Huascaran program and hence, they were assigned an innovation room coordinator, training and standard software. Additionally, the provision of internet access to these schools was prioritized.	Peru	Grades 7-11 (Grades 1-5 secondary school), 350 secondary schools.	18,049	No effects.	Not specified.
Cristia et al. (2017)	Program: "One Laptop per Child". Increased ratio of computers per student from 0.12 to 1.18 in treatment schools. 40-hour teacher training on how to use computers for pedagogical purposes. Randomization at school-level.	Peru	318 schools, 8 rural areas.	2,609	No effects.	USD 200 per laptop.
de Melo et al. (2014)	Program: “Plan Ceibal”. One computer per pupil, with data detailing time of delivery of computer to individual, therefore allowing to use a continuous treatment variable (days of exposure). Leveraging different delivery dates, researchers use variation in delivery date across individuals within same school with fixed-effects at individual and school-level.	Uruguay	Grades 3-6, 90 primary schools, nationally.	2,057	No effects in math and reading.	USD 180 per laptop.
Habyarimana and Jack (2018)	A mobile money platform operated a "lock savings account", especially targeted at parents about to incur high educational costs.	Kenya	Parents of children half way into grade 7 (final year of primary). Parents from 337 primary	4,020	Higher secondary school enrollment by 5-6 p.p. (ITT) or 18-24 p.p. (TOT). Total financial savings increased between three and four times. No effects on test scores.	Not specified, although the lock savings account earns a bonus 1% on top of the 2-5% APR (forfeited if

			schools in three counties.			savings are withdrawn beforehand).
Habyarimana and Sabarwal (2018)	Provision of eReaders. Testing the marginal effects of eReaders with instructional material from the pure effect of endowing the student with an eReader. Four experimental groups: a pure control group, a group that only received an eReader with only non-curriculum reading material, a group that received an eReader with non-curriculum material and curriculum textbooks, and a fourth group with all of these previous features, plus supplementary curriculum-relevant material. Randomization performed at the student-level.	Nigeria	Grade 8. Lagos; students came from 214 schools.	497	Overall no significant effects of eReader. Students that received eReaders with curriculum materials and no access to textbooks has large, imprecise effects. eReaders without curriculum material led to a decline in in overall reading and math.	Cost of eReader is USD 80.
Kho et al. (2018)	Impact evaluation of internet access on student performance in the universe of public primary schools in Peru that initially acquired internet between 2007 and 2014. Leverages variation in cohorts impacted, and timing of rollout to schools.	Peru	Grade 2 provides test scores, but policy affected Grades 1-6. 5,903 public primary schools.	218,883	Initial math improvements of 0.042-0.076 SD, growing at a rate of 0.047 SD per year, reaching 0.29 SD 5 years after installation.	Not specified.
Malamud and Pop-Eleches (2011)	Program: “Euro 200 Program”. USD 300 Voucher only valid to buy a home computer. Educational software needed to be installed separately, not always installed. Teacher training, 530 multimedia lessons on the use of computers for educational training.	Romania	Grades 1-12, Between 25,051 and 35,484 families received vouchers of program yearly between 2004 and 2008.	3,354	-0.44 SD math GPA, -0.56 SD in Romanian, -0.63 SD in English, higher scores in computer skills test by about 0.33 SD.	USD 300 per voucher plus management cost (not specified).
Malamud et al. (2019)	Three experimental arms: students that received computers with access to high-speed internet, students that received computers without access to high-speed internet, and a pure control group. Lotteries to give away 4 laptops within each class. Computers had standard software and some educational games. Randomization at student-level within classes in treatment schools.	Peru	Grades 3-5, 14 low-achieving public primary schools.	2,126	No effects in learning, cognitive and noncognitive skills. Free internet access led to improved computer and internet proficiency.	Not specified.
Mensch and Haberland (2018)	Program: GirlsRead! Three experimental branches: a pure control branch, a second branch with safe spaces for girls where mentors facilitate an empowerment-based life-skills curriculum and all the activities of the second branch, plus e-readers that girls keep for the duration of the program with approximately 100 books of varying reading levels primarily written by African authors. Randomization at school-level.	Zambia	Grade 6. 36 schools in three districts.	1,299	Reading scores 4.6 p.p. higher in e-reader arm. Three quarter of girls attended all community sessions. Only 2.4% of all e-readers were lost, stolen, or broken.	Not specified.
Meza-Cordero (2017)	Impact evaluation of One-Laptop-per-Child" intervention, using a difference-in-differences estimation strategy, as treatment was not randomly assigned.	Costa Rica	Grades 1-6. 34 schools.	3,174	Increase in time using a computer (to browse internet, do homework, read, and play), decrease of time spent doing homework and outdoor activities; no effects on learning.	USD 225 per student accounting for all costs, USD 209 per computer.
Mo et al. (2013)	Evaluation of One Laptop per Child policy. Randomization at individual-level.	China	Grade 3. 13 schools of migrant children in Beijing.	300	Effects in computer skills of 0.32 SD, 0.17 SD in math, no effects on language.	Not specified.
Navarro-Sola (2019)	Program: Expansion of Mexican Telesecundaria, or schools using televised lessons. The study exploits the staggered rollout of the policy from 1968 to present.	Mexico	Grades 7-9, 3,132 telesecundarias in 2,110 localities.	896,274	For every telesecundaria per 50 children, 10 more children enroll in secondary education, and 2 more pursue further education. Every year of education induced by telesecundaria, increased income by 17.6%.	USD 704 per student per year, including all administrative costs.

Piper et al. (2016)	Four experimental groups: base PRIMR program (early literacy program focused on teacher training, instructional support, and student learning materials at 1:1 ratio), PRIMR plus a tablet for the teacher to scaffold their instruction, PRIMR for pupils e-readers with age-appropriate textbooks, and a control group. Although there was randomization at the school-level, there were still imbalances in baseline characteristics, so authors prefer a difference-in-differences strategy.	Kenya	Grade 2. 80 schools in Kisumu county.	1,580	All treatment arms had positive effects ranging from 0.17-0.29 SD in English, and 0.26-0.32 SD in Kiswahili. The most effect arm was the basic PRIMR arm.	Cost of tablet is USD 150, cost of eReader is USD 70. The cost of the basic PRIMR program was USD 2.28 per pupil per subject per year.
Pitchford (2015)	Three experimental arms: math tablet intervention, non-math tablet control, and standard face-to-face practice. Intervention lasted 8 weeks, for 30-min per day. The math tablet intervention consisted of four different apps developed by onebillion©. Apps based on the National Primary Curriculum Randomization at individual level.	Malawi	Grades 1-3, One medium-sized urban primary school.	318	Positive, and statistically significant effects in math and language. Authors do not provide enough information to translate into SD units.	Not specified.
Seo (2017)	Program: GivePower school program. Six experimental groups: G1 schools received two 0.12 kWh solar home systems including lights and TVs ("facilities"); G2, solar facilities and English videos; G3, solar facilities and bilingual videos; G4, English videos only; G5, bilingual videos only; and control schools.	Tanzania	Grade 11, 164 schools in northern Tanzania. Schools are between the national median (57) and the mean (75) in terms of enrollment.	11,697	Impact of solar-facilities-enabled programs, averaged across video-provision status, to be 0.05 SD on secondary exit exam (across all subjects), and 2.8 p.p on passing rates.	USD 6.41 per student.
Yanguas (2020)	Analysis of long-term effects of "Plan Ceibal", or a one-laptop-per-child in Uruguay (whose short-term results are described in de Melo, et al., 2014). Study leverages cross-cohort variation and it is the first study with long-term, causal estimates of this kind of policy.	Uruguay	Adults exposed to one-laptop-per-child policy as children. All students in public primary and middle schools .	12,775	No effects on educational attainment as an adult. For college-goers, enrollment in the program led to lower likelihood of enrolling in science and technology majors.	Same as in de Melo (2014), et al. USD 180 per laptop.

Notes: All randomized controlled trials indicate the level at which units were randomized. For the full coding and more detailed information on all the core studies included in the review, please see this online document. The statistical significance of the findings stems from what each of the studies reports, and the alpha threshold for significance may vary by disciplinary approach of each paper. Abbreviations: "p.p.": percentage points, "SD": standard deviations.

Online Table 2: summary of studies included within the "Technology-enabled behavioral interventions" category

Study	Intervention	Context	Target grade and outcomes	Sample	Findings	Cost
Adelman et al. (2015)	Directors received a smartphone with a built-in system to allow school directors to send information about the school to a centralized server, including daily photographs of teachers to verify presence. School inspectors could then access the server in real time for efficient supervision.	Haiti	Teachers. 200 public and private primary schools.	2,260	No effects on test scores. The program did not improve management practices such as record keeping either. Low take-up.	Not specified.
Aker and Ksoll (2019)	Treatment consisted of a mobile phone monitoring program, where students, teachers and village chiefs were called on a weekly basis, over a six-week period. No phones or incentives were provided. 140 schools were assigned to an adult education program, and 20 to the pure control group. Among the 140 schools, half were assigned to monitoring. Randomization at village-level.	Niger	Adult learners. 160 villages, stratified by regional, and sub-regional administrative divisions.	1,776 individuals, 160 villages.	Monitoring increased reading by 0.14-0.30 SD, and math by 0.08-0.15 SD. Villages with no monitoring had no effects relative to the pure control villages.	Overall reported cost of mobile monitoring was USD 6.5 per village.
Berkhout et al. (2020)	Impact evaluation of the effect on test scores (implicitly on cheating) of switching to computer-based testing (CBT) for the high-stakes, national examination of junior secondary schools in Indonesia.	Indonesia	Grades 9 and 12. 50,124 junior secondary schools nationally.	353,190 students.	The introduction of computer-based testing (CBT) decreased scores by 0.40 SD, interpreted as a decrease in cheating. However, results become insignificant after two years of the introduction of CBT, suggesting that actual learning had to happen to compensate for the loss in test scores due to the curtailing of cheating.	Not specified.
Berlinski et al. (2016)	Program: "Parents up to date". High-frequency information about their selected child via text message (SMS messages). SMS texts contained specific information on attendance, behavior, and math test scores of each parent's child. Randomization at individual-level, along with share of students treated in each class.	Chile	Grades 4-8. 85 classes in metropolitan area.	1,447	0.08 SD in math after only 4 months. Probability of passing a grade increased by 2.8 percentage points. Increase probability of attending school for more than 85% of the time (threshold needed for grade progression) by more than 6.6 p.p.	Not specified. "Low-cost intervention".
Duflo et al. (2012)	Teacher attendance in treatment schools was monitored using cameras, and their salaries was linked to their attendance. Instructions for one student to take a picture of the teacher at the start and end of the work day. Cameras has tamper-proof date and time functions. Attendance was tracked for 30 months. Randomization at school-level.	India	Teachers. 113 single-teacher non-formal education centers/schools in rural villages of Rajasthan.	113 teachers. 2,230 students at baseline.	Teacher absenteeism fell by 21 percentage points, and test scores increased by 0.17 SD.	Not specified.
Dustan et al. (2019)	SMS campaign to increase civil servants' compliance with maintenance activities. Each SMS contains a fixed and a variable component. The fixed component includes the bureaucrat's first name and the deadline for task compliance. These fixed elements are rooted in behavioral insights. The variable component is the main behavioral lever, which could be a reminder/warning, social norm, monitoring, shaming, auditing threat, or a control condition. Randomization at school-level.	Peru	Civil servants in charge of a school maintenance program. 24,000 schools across Peru.	24,268	Increase of 3.86 p.p. in the probability of submitting an expense report by deadline, no evidence that the SMS campaign affected the quality of most of the infrastructure items.	Total cost of 57,860 SMS was USD 1,273, and the labor costs associated with the programming and sending of the SMS were USD 188 for the full campaign.
Gaduh et al. (2020)	Intervention had three different treatment arms. The first treatment arm provides a scorecard which evaluates the use of a government	Indonesia	Teachers, 270 mostly public schools in 5 districts.	3,832 students, 827 teachers.	Gains across all treatment arms; largest in treatment arm with camera: 0.18 SD in language, 0.20 SD in Math.	USD 40 per student.

	allowance. The second and third treatments added to the first treatment a pay-for-performance scheme that relied on included the first treatment. The second treatment added a camera with a timestamp which made the allowance dependent on teacher presence. The third treatment the payment of the allowance depended on the result of the scorecard. Randomization at school-level.				Camera treatment arm showed positive, imprecise estimates on teacher behavior, working at school, and teaching in class.	
John et al. (2016)	Impact evaluation of an electronic career guidance package for secondary schools, the e-Career Guidance System.	Nigeria	Grade 8, 2 public secondary schools in Akwa Ibom state.	60 students.	Positive effects in vocational and career outcomes. Not enough information to translate gains into SD units.	Not specified.
Neilson et al. (2018a)	Videos and infographics informing about the returns to education at different educational levels. Randomization at the school-level.	Peru	Grades 1-11, but learning outcomes only measured for Grade 8. 2,626 public schools in all department capitals across Peru, and 250 rural schools in Cusco and Arequipa.	Not specified.	Reduction of school dropout in urban areas (after second year of implementation, once take-up of treatment was higher) by 1.8 p.p., or 18.8% of the baseline; in rural areas the reduction was 7.2 p.p. or 50% of the baseline. Effects on math were 0.04 SD, and on reading were 0.03 SD.	At the scale of 25,000 students, authors estimate the cost would be USD 0.06 per student.
Neilson et al. (2018b)	Videos and infographics informing about the returns to education at different educational levels. Randomization was at the school-level, where 1524 schools were selected for treatment.	Dominican Republic	Grades 7-12, 2,469 public schools.	~120,000	Preliminary results show that the informative and persuasive videos both led to decreases in school dropout, and increases in standardized test scores.	Major costs were production and elaboration of the videos ($104,000).
Riley (2017)	Students watched a film projection of "Queen of Katwe", a movie about a teenage girl from the slums of Kampala, Uganda striving to become a chess master, as a way to change students' beliefs about the importance of education. Randomization at the student-level.	Uganda	Grades 10 and 12, 13 secondary schools in urban Kampala.	1,446	0.11 SD in math for grade 10, 0.13 SD in math for grade 12; 9 percentage points more likely to continue enrolled in secondary school.	USD 5 per student.
Vakis and Farfan (2018)	SMS campaign with potentially useful information for teachers, such as reminders about deadlines, teacher benefits, motivational texts, and occupational wellness. No pure control group, as control group got at least two informative texts, and once on Teachers' day. The teacher's name was in some messages.	Peru	Teachers. 35,000 schools nationally, only teachers that registered for the program.	Experimental sample: 13145 teachers, rolled out nationally to 186,000 teachers.	3 p.p. increase in questions about job satisfaction and motivation. Likely underestimate, given that there was no pure control group.	Each SMS costs USD 0.03.

Notes: All randomized controlled trials indicate the level at which units were randomized. For the full coding and more detailed information on all the core studies included in the review, please see this online document. The statistical significance of the findings stems from what each of the studies reports, and the alpha threshold for significance may vary by disciplinary approach of each paper. Abbreviations: "p.p.": percentage points, "SD": standard deviations.

Online Table 3: summary of studies included within the "Improvements to instruction" category

Study	Intervention	Context	Target grade and outcomes	Sample	Findings	Cost
Angrist et al. (2020)	Two low-technology interventions to substitute schooling during this period: SMS text messages with "problems of the week", and direct phone calls with intruction (15-20 minutes) plus the SMS. Randomization at the student-level.	Botswana	Grades 3-5. 103 schools across 9 out of 10 regions in Botswana.	4,550	0.16 SD in math from SMS intervention, and 0.29 SD in math from phone call intervention. Increased parental engagement.	USD 2.13 per child for only SMS intervention, and USD 14 per child in the phone and SMS intervention.
Beg et al. (2019)	Program: “eLearn”. Program delivers expert math and science content through short videos with multimedia presentations, for four months of exposure. Curriculum tailored to local 8th grade curriculum. After each lecture, there would be multiple-choice review questions, a small tablet for teachers to project the material for their own review, and an LED screen installed in each classroom. Some teacher training on how to use the tablets was provided. 29 hours of content during regular class time. Randomization at school-level.	Pakistan	Grade 8. 100 schools in Punjab.	2,622	0.26 SD in Math, 0.26 in Science, 0.33 SD in combined score. Small increases in student and teacher attendance.	USD 15 per student with the inclusion of high fixed-costs at the scale of 100 schools, USD 9 was the marginal cost per student.
Berlinski and Busso (2017)	Program: testing a pedagogical intervention designed to give students a more active role in learning geometry, along with different technological complements. One pure control group and four treatment arms: 1) active learning, 2), active learning plus an interactive whiteboard, 3) active learning plus a computer lab, 4) active learning plus one computer per student. Randomization at the school-level.	Costa Rica	Grade 7. 85 schools.	18,000 students and 190 teachers. Sample was nationally representative.	Negative effects of -0.17 SD for active learning alone, and -0.25 SD for active learning plus technology. No treatment arm had positive effects. High take-up by teachers.	Not specified.
Bianchi et al. (2019)	Evaluation of government reform that connected high-quality teachers in urban areas with more than 100 million students in rural middle schools through satellite internet over four years. First difference in cohort, and second difference in geographic location, leveraging staggered implementation.	China	Middle schoolers, Rural schools in China.	4,479	0.18 SD in math 7-10 years later, 0.21 SD in Chinese. Share of people investing in informal education increased 9.8 p.p., earnings increased, increased likelihood of being in more analytical and less manual jobs, increased internet and computer usage.	Project served 100 million students, costing CNY 8.78 billion (USD 1.24 billion), or approximately USD 12.4 per student served.
Blimpo et al. (2020)	The program targets math and science instruction through incorporation of technology that enhances students’ participation. The program provided computers for teachers, scripted lessons, and customized software; equipped classrooms with smart projectors (smartboards) and handheld devices (smart responders) that students can use to respond to teachers; as well as provided textbooks for students. Treatment also included “student responders”, are battery-operated, wireless handheld devices that allow students to provide responses simultaneously, and allows teachers to monitor and track students’ responses in real-time.	Gambia	Grade 12 (measured outcomes), program for grades 1-12, 24 schools across the Gambia	1044	0.54 SD on Math, 0.20 SD in English, increased probability of passing secondary exit exam by 15 p.p.	~USD 3,000 per classroom.
Böhmer et al. (2014)	After-school mathematics intervention aimed to fill knowledge gaps using computer-assisted learning (CAL). Khan Academy resources were used to teach basic numeracy. Each individual has full autonomy over which exercises they attempt. Gamefication is used to incentivize and engage the learners Randomization at the individual-level.	South Africa	Grade 8. 9 schools in Western Cape circuit, which had to meet the criteria of good management and a working computer laboratory with an internet connection.	472	0.32 SD on basic numeracy questions, and 0.25 SD on core grade 8 curriculum questions.	Not specified.
Borzekowski (2018)	Showing of educational videos at school, part of the "Akili and Me"	Tanzania	Pre-school. 9 randomly selected	595	Positive effects across several fields of basic	Not specified.

	series. "Aliki and Me" is am animated series teaching school readiness skills, in both Kiswahili and English. The videos were contextually-relevant and sensitive. Randomization at the student-level.		schools in peri-urban areas of Morogoro.		numeracy and literacy. ~0.15 SD in English and 0.22 SD in counting.	
Borzekowski and Henry (2010)	Showing of "Jalan Sesama", a multimedia educational project, developed for Indonesian children. Television episodes presenting educational messages regarding literacy and numeracy, health and safety, social development, and environmental and cultural awareness. Randomization at the individual-level.	Indonesia	Children age 3-6. Children selected from remote areas which typically have poor reception of broadcast television three main locations (Munjul, Kota Dukuh, and Gunung Batu village) from the Munjul subdistrict.	160	0.12 SD fin early cognitive skills or the low-exposure group and 0.35 SD for the high-exposure group.	Not specified.
Borzekowski et al. (2019a)	Evaluation of the adaption and testing the Tanzanian-made program, Akili and Me (studied in Borzekowski), for children's viewing in Rwanda. Randomization at the student-level.	Rwanda	Pre-school to grade 2. Randomly-selected kindergartens and primary school in Gihara.	434	Statistically significant increases in math and language. Not enough information provided to reliably convert coefficients into SD units.	Not specified.
Borzekowski et al. (2019b)	Showing of Galli Galli Sim Sim, the Indian version of Sesame Street, 30 min of television five days a week for twelve weeks, varying how much Galli Galli Sim Sim versus other programming children watched. Randomization at the school-level.	India	Pre-school, 99 preschools in Lucknow, with children ages 3-7.	1.340	Overall literacy score reports effects between 0.24-0.37 SD, and numeracy scores effects of 0.15-0.20 SD.	Not specified.
Cilliers et al. (2020)	Three year follow up of Kotze et al. (2019).	South Africa	Grades 1-3, 180 public schools located in low-income rural communities in the Mpumalanga province.	2,684	After 3 years, the in-person coaching arm achieved improvements in oral language of 0.31 SD and reading proficiency of 0.13 SD. The in-person treatment arm achieved gains in oral language of 0.12 SD and no gains in reading proficiency. Furthermore, the virtual coaching induced a negative effect on home language literacy.	The cost per learner per year of the on-site program is USD 66, and the cost per learner per year of the virtual program was USD 51.
De Hoop et al. (2020a)	Evaluation of a "e-School 360" model, a multi-faceted program that integrates technology into education, provides ongoing teacher training and professional development, and includes community ownership.	Zambia	Grades 1-3. 64 schools across 3 rural districts in the Zambia's Eastern Province.	1,924	0.33 SD in reading, and 0.14 SD in math.	The cost of the program was USD 3 per month per student.
Gambari et al. (2016)	Video-based cooperative, competitive and individualized instructional strategies on the performance of senior secondary schools' students in geometry in Nigeria. The treatment involved identification of some difficult concepts in mathematics that were developed in simpler instructional module using video instruction platform. Randomization at the school-level.	Nigeria	Senior secondary students, 4 secondary schools in Minna.	120	Positive effects on all treatment arms, not enough information to translate into SD units.	Not specified.
Johnston and Ksoll (2017)	Broadcasting live instruction via satellite to rural primary school students. Classrooms in 70 randomly selected schools equipped with the technology required to connect to a studio in Accra. Randomization at school-level.	Ghana	Grades 2-4, 144 schools, districts of the Volta and Greater Accra regions; districts classified by Ghanaian government as "deprived".	4,545	0.23 SD in math, no effects in reading fluency overall, but gains in foundational skills (letter and word identification), no effects on classroom attendance nor time-on-task.	USD 22 per student, as authors estimate USD 100 per standard deviation gained. Estimate includes fixed-costs, which authors claim to be a large proportion of total costs.
Kotze et al. (2019)	Two different versions of coaching within a structured pedagogic program, the conventional form of one-on-one	South Africa	Grades 1-3. 180 public schools located in low-	3,227	Not enough information to convert point estimates into SD units. However,	The per-student costs of the on-site coaching and the

	on-site instructional coaching, and virtual coaching, which involves using a tablet, cellular phone calls, and daily text messaging.		income rural communities in the Mpumalanga province.		researchers find that "students from the two intervention groups performed consistently better than the control students" on most numeracy and literacy tasks.	virtual coaching models do not differ dramatically, and are US$48 and US$43, respectively, per year.
Lehrer et al. (2019)	Evaluation of "Sankoré" equipment, which consisted of classroom provision of interactive whiteboards.	Senegal	Grades 1-2. 173 initially surveyed schools in Dakar, Diourbel, Kaolack, Thiès, and Fatick.	14,713	0.186 SD in math.	Not specified.
Näslund-Hadley et al. (2014)	Program: "Tikichuela". Intervention consists of interactive audio segments that cover the entire preschool math curriculum. Since Paraguayan classrooms tend to be bilingual, the audio and written materials use a combination of Spanish and Guaraní. Audio lessons were implemented four days a week, with one day set aside to review what had been learned during the week. This extra day gave teachers flexibility to review topics that, according to their observation, the children needed more practice or assistance in addressing. The average duration of each class was 60 minutes. Randomization at the school-level.	Paraguay	Pre-school. 265 schools in department of Cordillera	2,907	0.16 SD in math.	Not specified.
Naik et al. (2016)	Technology-assisted teaching to replace one-third of in-school instructional time. Intervention combines computers and broadband connectivity with more conventional satellite technology to deliver classes taught by expert teachers at a central location using multimedia teaching aids. These lectures cover the standard syllabus prescribed for all schools in the state by the State Department of Education.	India	Grades 5-10. 1,823 rural, public schools across 18 districts in Karnataka. Data collection performed only in sub-sample of 105 treatment schools, and 98 comparison schools.	14,084	0.1-0.2 SD in math, 0.2-0.3 SD in science, 0.2-0.4 in English.	USD 1.7 per student per year.
Wennersten et al. (2015)	Program: BridgeIT. Teachers of Standard 5 and 6 English and Science classes were notified of the availability of new videos via text messages (SMS), which they downloaded onto their phones using an open-source application and showed, with suggested activities, to students on a TV screen using a TV-out cable. Participation was not randomized, it was simply rolled out in certain schools first, chosen by funders and implementers.	India	Grades 5 and 6, 86 schools in Andhra Pradesh and Tamil Nadu.	3,327	0.36 SD in English in both states. 0.98 in Science in Andhra Pradesh. Science gains not reported for TN.	USD 10.50 per student.
Wolf et al. (2018)	Three experimental arms: teacher training, teacher training plus parental-awareness meetings, and controls. The programs incorporated workshops and in-classroom coaching for teachers. The technology portion was the video-based discussion groups for parents. Randomization at the school-level.	Ghana	Teachers in public and private kindergartens in the Greater Accra Region, 240 schools.	444 teachers, and 3345 children.	Treatment arm with parental intervention has effects of ~0.14 SD in overall school readiness, ~0.09 SD in math, ~0.08 in literacy. The branch without parental intervention had slightly higher, statistically significant effects. Parental meetings had no effect no the effectiveness of the teacher training.	Not specified.

Notes: All randomized controlled trials indicate the level at which units were randomized. For the full coding and more detailed information on all the core studies included in the review, please see this online document. The statistical significance of the findings stems from what each of the studies reports, and the alpha threshold for significance may vary by disciplinary approach of each paper. Abbreviations: "p.p.": percentage points, "SD": standard deviations.

Online Table 4: summary of all studies included within the "Self-led learning" category

Study	Intervention	Context	Target grade and outcomes	Sample	Findings	Cost
Abrami et al. (2016)	Interactive, multimedia literacy software for 90 minutes per week, for 13 weeks. Randomization at the class-level.	Kenya	Grade 2. 12 classes.	429	Gains in certain areas such as reading comprehension and listening skills. Not enough condensed information to translate into gains in SD.	Not specified.
Araya et al. (2019)	Program: "ConectaIdeas", two weekly, 90-minute sessions in a computer lab where students solve math exercises. Software can create individual and group competitions. Competitions were intra- and inter-schools. Software shows each student how many exercises have been completed, and compares it with class average. Personalized "ads" are shown regularly to motivate students. Randomization at the class-level.	Chile	Grades 4, in 24 schools. Public schools in Chile attended by socioeconomically disadvantaged students who also significantly lagged in math achievement.	1,089	0.27 SD in math, no effect in language. Increased students' preference to use technology for math learning, promoted the idea that studying can raise intelligence. Increased math anxiety and reduced willingness to collaborate in groups.	USD 150 per student cost, 5% increase in public expenditure per primary student in Chile
Bai et al. (2016)	Computer-assisted complement to English class. Comparison between "computed assisted instruction" (CAI; program integrated with curriculum), "computer assisted learning" (CAL; not integrated into teacher's instruction), and a pure control group. The integrated program included three parts: a curriculum, a lesson-by-lesson English Teaching Plan, and a set of instructions on teacher responsibilities. English teachers in CAL and CAI were also compensated with 80 USD per semester. Randomization at school-level.	China	Grade 5 in 127 schools. Rural schools in Haidong Prefecture in Qinghai Province.	6,304	No effects of pooled test for CAI/CAL, effects of 0.07 SD for CAI when tested separately. Suggestive evidence that CAL did help higher performers.	Not specified.
Banerjee et al. (2007)	Program: Pratham-developed program during year 1, program developed by Media-Pro during year 2. Two hours per week during or before/after school, with two children per computer. Software linked to Gujarat's curriculum, focusing on basic skills. Software changes the question difficulty by ability. Randomization at the school-level.	India	Grade 4. 110 schools. Mumbai and Vadodara.	~5,500	0.35 SD in math for year 1; 0.48 SD in math for year 2. Math effects persisted one year after leaving intervention. No effect on language either year. No effect on attendance.	USD 15 per student per year.
Bettinger et al. (2020)	Intervention tested computer-assisted learning program, with theoretical implications for estimation of educational production function. Three treatment arms: a base dosage CAL arm with ~20-25 minutes per week of math CAL and ~20-25 minutes of language CAL; a double-dosage CAL arm with ~40-50 minutes of math CAL and ~40-50 minutes of language CAL; and a control arm. The software is adaptive to each student's level. Randomization at the class-level.	Russia	Grade 3. 343 schools.	5,621	0.11-0.12 SD in math for base dosage, and similar results for the double-dosage-level arm. 0.06-0.07 in language for the base dosage arm, and no effects in language for the double-dosage arm. The differences between the two treatment arms are not statistically significant in either subject.	Not specified.
Brown et al. (2020)	Program evaluation of a digital game-based learning program ("Can't Wait to Learn").	Sudan	Children age 7-9. 8 villages in Sudan	221	Positive effects in math, Arabic, and psychological well-being. Not enough information provided to translate gains into SD units.	Not specified.
Büchel et al. (2020)	Comparison of relative effectiveness of computer-assisted learning (CAL) and traditional teaching. The first treatment arm is pure CAL, the second is CAL plus traditional teaching, and the control group is traditional classroom teaching. Each experimental arm consisted of 90 minutes of additional instruction per week. Randomization at the school class-level.	El Salvador	Grades 3-6. 198 school classes in Morazán across 29 schools.	3,197	0.21 SD from CAL, 0.24 SD of CAL plus supervisor (difference not statistically significant).	The cost per child is 44 USD for the traditional teaching arm, 43 USD for the CAL arm, and 56 USD for the CAL plus teacher.

Carrillo et al. (2011)	Program: "Personalized Complementary and Interconnected Learning (APCI) program". Computer-aided instruction in mathematics and language, 3 hours per week during school. Personalized curriculum based on screening test; fixed after screening test. Randomization at the school-level.	Ecuador	Grade 5. 16 schools. Public schools in Guayaquil.	1,061	0.30 SD in math, and no effect on language. Larger gains for students at the top of the achievement distribution.	Not specified.
Chong et al. (2020)	Mandatory six-month Internet-based sexual education course. Randomization at the school*classroom level (to allow for analyses of spillovers).	Colombia	Grades 9. 138 classes across 69 junior high schools in 21 Colombian cities.	4,599	0.4 SD increase in knowledge about sexual education, 0.2 SD in attitudes, and 55% increase in likelihood of redeeming vouchers for condoms.	USD 14.7 per student per semester.
De Hoop et al. (2020b)	Program evaluation of a digital game-based learning program ("Can't Wait to Learn").	Jordan	Grades 1-3. 35 schools within Zarqa Governorate.	709	No effects in math, Arabic, and psychological well-being.	Not specified.
Derksen et al. (2020)	Evaluation of program providing access to Wikipedia. Randomization at the student-level.	Malawi	Grades 8-10. 4 government boarding schools.	1508	Gains in English of 0.103 SD.	USD 4 per student.
Freeman and Hawkins (2017)	Evaluation of "Evoke", a game-based interactive environment. Evoke is a project-based learning module, using storytelling, virtual games, and social networks, which connects students with their peers and mentors. Randomization at the class-level.	Colombia	University students, two thirds being between 18-22 years old. Recruitment in 14 university classes.	297	Gains in "21st century and socioemotional skills". Authors do not provide enough information to translate gains into standard deviation units.	Not specified.
He et al. (2008)	Two interventions, only one of which involves an EdTech intervention. This intervention consists of a "PicTalk" machine, which is designed to be used by a single student who with the help of a stylus, can point to pictures and hear the word pronounced. Learner could choose topics, and within each topic, what words to point to. The other, non-EdTech, intervention consisted of sets of flashcards designed to cover the same competencies as the PicTalk machine. Randomization at the school-level.	India	Grades 1-5. 97 schools in Thane Municipal School District, and 242 schools in Mangaon sub-district government schools.	15,062 students across all years, all schools.	0.25-0.35 SD, depending on specification. Stronger students benefit more from the more self-paced machine-based implementation.	USD 20.46 per student in Thane, and USD 11.20 per student in Mangaon (including costs of machines and material development).
Hirshleifer. (2016)	Treatment consists of a math software curriculum implemented in all classrooms of the intervention. The main research question focuses on whether incentivizing inputs (the completion of learning modules) is more effective than the incentivizing of outputs (a test at the end of each module). The incentives were small monetary rewards. Randomization at the treatment level using a partial rotation design.	India	Grades 4-6. 45 classrooms in Mumbai and Pune.	3,218	0.57 SD in math for the branch incentivizing the inputs, and 0.24 SD for the branch incentivizing outputs.	Maximum incentive was USD 2.65 per student (200 rupees of rewards).
Ito et al. (2019)	Treatment consisted of 20 30-minute classes when students were allowed to use an app-based computer-aided instruction instead of regular math classes. Adaptive learning with algorithm in response to the proficiency level of each individual. Randomization was at the class-level.	Cambodia	Grades 1-4. 5 public elementary schools near Phnom Penn.	1,636	0.56-0.67 SD in math scores, increases in subjective expectation of being able to attend tertiary education. No effects on motivation.	Not specified.
Jere-Folotiya et al. (2014)	Evaluation of computer-based literacy game. Randomization at the student-level.	Zambia	Grade 1. 42 government schools in Lusaka.	573	Positive effects in spelling. Not enough information to translate into SD units.	Not specified.
Lai et al. (2013)	Two 40-min mandatory sessions per week during lunch breaks or after school, teams of 2 children. Based on national curriculum. Reinforced material taught that week Program was remedial in nature. Randomization at the school-level.	China	Grade 3 and 5. 72 schools rural boarding schools in Shaanxi.	2,726	0.12 SD in math, no effects in language across both grades.	Not specified.
Lai et al. (2015)	Two 40-min mandatory sessions per week during lunch breaks or after school, teams of 2 children. Based on	China	Grade 3. 43 migrant schools in Beijing.	2,369	None in language, 0.15 SD in math, 0.31 points in 1-10	Not specified.

	national curriculum. Reinforced material taught that week Program was remedial in nature. Randomization at the school-level.				scale asking about whether child "likes school".	
Lai et al. (2016)	Two 40-min mandatory sessions per week during lunch breaks or after school, teams of 2 children. Based on national curriculum. Reinforced material taught that week Program was remedial in nature. Randomization at the school-level.	China	Grade 3. 57 rural schools in Qinghai.	6,865	0.15 SD in both math and language.	USD 7.6 per student.
Linden (2008)	Program: Gyan Shala Computer Assisted Learning program. Two children with one computer (split screen), two versions of the treatment. Version 1: one hour per during school, version 2: one hour per day after schools. Reinforces material taught that day. Randomization at the school-level.	India	Grades 2-3, 60 schools. Gyan Shala schools in Gujarat.	779	-0.57 SD in math as a substitute, and 0.28 SD in math as a complement.	USD 5.2 per student.
Lysenko et al. (2019)	Evaluation of computer-based literacy game.	Kenya	Grades 1-3. 48 classes	1,899	Positive effects in spelling. Not enough information to translate into SD units.	Not specified.
Ma et al. (2020)	Three experimental branches: 1) pure control group, 2) supplemental computer-assisted learning, 3) supplemental workbook. The program sessions were held once a week for 9 months. Randomization happened at the class-level.	China	Grades 4-6. 130 schools from 9 impoverish counties.	4,024	No effects of the pure technology portion of the intervention.	USD 18 per student.
Mo et al. (2014a)	Two 40-min mandatory sessions per week during lunch breaks or after school, teams of 2 children. Based on national curriculum. Reinforced material taught that week Program was remedial in nature. Randomization at the school-level.	China	Grade 3, and 5. 72 rural schools in Shaanxi.	4,757	0.17 SD in math.	USD 9,439 in total over one year.
Mo et al. (2014b)	Two 40-min mandatory sessions per week during lunch breaks or after school, teams of 2 children. Based on national curriculum. Reinforced material taught that week Program was remedial in nature. Randomization at the school-level.	China	Grade 3, and 5. 72 rural schools in Shaanxi.	2,741	0.25-0.26 SD in math.	USD 9,439 in total over one year.
Muralidharan et al. (2019)	Program: "Mindspark". Evaluation of after-school Mindspark centers, which scheduled 6 days of instruction per week, with 90 minutes per day, for 4.5 months. Half of each session was self-driven learning on Mindspark software, and the other half consisted of instructional support from a teaching assistant in groups of 12-15 students. Technology-led instructional program, software benchmarks the initial learning level of every student and dynamically personalize the material to match the level and rate of progress made by each student. Randomization at the student-level.	India	Grades 4-9. Students recruited from 5 public middle schools in Delhi.	619	0.37 SD in Math, 0.23 in Hindi.	USD 15 per student per month.
Pitchford et al. (2018)	Three experiments reported, testing the effectiveness of apps developed by onebillion©. Eighteen 30-min sessions on average across the 14-month study period. Note that treatment was not randomly selected, but rather the government chose one school per district to be treated, and researchers chose a similar comparison school. Hence, this is closer to PSM than to an RCT.	Malawi	Grades 1-2. 14 schools across seven education districts across Malawi.	1,217	Gains in math in the order of 0.19-0.62, depending on gender, and gains of 0.33-0.46 in reading. Girls benefited more from the intervention.	Not specified.
Rosas et al. (2002)	Introduction of educational video-games in the classroom. Students in the experimental group were exposed to an average of 30 hours over a three-month	Chile	Grades 1-2. Economically disadvantaged schools.	1,274	Positive, and statistically significant effects in math and language. Authors do not provide enough	Not specified. "Low-cost videogame".

	period. The games had a self-regulation system that dynamically adapted the level of difficulty of the contents to the player's learning pace, presenting the player contents based on his or her level of knowledge. The games had a progressive and increasing level of difficulty, based on the presentation of antagonists and obstacles. According to the child's performance, the game provided feedback indicating if he or she chose the correct or incorrect answer.				information to translate into SD units.	

Notes: All randomized controlled trials indicate the level at which units were randomized. For the full coding and more detailed information on all the core studies included in the review, please see this online document. The statistical significance of the findings stems from what each of the studies reports, and the alpha threshold for significance may vary by disciplinary approach of each paper. Abbreviations: "p.p.": percentage points, "SD": standard deviations.