

Girls' Education at Scale

David K. Evans, Amina Mendez Acosta, Fei Yuan

Abstract

Many educational interventions boost outcomes for girls in settings where girls face educational disadvantages, but which of those interventions are proven to function effectively at large scale? In contrast to earlier reviews, this review focuses on large-scale programs and policies—those that reach at least 10,000 students—and on final school outcomes such as completion and student learning rather than intermediate school outcomes such as enrollment and attendance. Programs and policies that have boosted school completion or learning at scale across multiple countries include school fee elimination, school meals, making schools more physically accessible, and improving the quality of pedagogy. Other interventions, such as providing better sanitation facilities or safe spaces for girls, show promising results but either have limited evidence across settings or focus on intermediate educational outcomes (such as enrollment) or post-educational outcomes (such as income earning) in their evaluations. These and other areas with limited or no evidence demonstrate many opportunities for education leaders, partners, and researchers to continue innovating and testing programs at scale. We discuss three considerations for incorporating evidence-based solutions into local education policies—constraints to girls' education, potential solutions, and program costs—as well as lessons for scaling programs effectively.

JEL codes: I21; I24; J16; O15

Keywords: education; gender; girls' education; inequality

Girls' Education at Scale

David K. Evans
Center for Global Development

Amina Mendez Acosta
Center for Global Development

Fei Yuan
Harvard University

This paper was made possible by financial support from Co-Impact and the Bill & Melinda Gates Foundation. The authors thank Varja Lipovsek and others on the Co-Impact team for guidance and Radhika Bula for sharing studies from the J-PAL post-primary education initiative. Kathleen Beegle, Christine Beggs, Erin Ganju, Nicole Haberland, Pamela Jakiela, Owen Ozier, and Dana Schmidt provided useful comments. Author names are listed alphabetically. Corresponding author: Evans, devans@cgdev.org.

David K. Evans, Amina Mendez Acosta, Fei Yuan. 2021. "Girls' Education at Scale" CGD Working Paper 594. Washington, DC: Center for Global Development. <https://www.cgdev.org/publication/girls-education-scale>

Center for Global Development
2055 L Street NW
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

The Center for Global Development works to reduce global poverty and improve lives through innovative economic research that drives better policy and practice by the world's top decision makers. Use and dissemination of this Working Paper is encouraged; however, reproduced copies may not be used for commercial purposes. Further usage is permitted under the terms of the Creative Commons License.

The views expressed in CGD Working Papers are those of the authors and should not be attributed to the board of directors, funders of the Center for Global Development, or the authors' respective organizations.

Contents

Introduction	1
The challenges	2
The broad policy challenge	2
The knowledge challenge	4
The methods used for this review	7
The solutions	9
Make school cheaper	9
Make school more physically accessible	11
Teach better	12
Gender focused interventions	15
Other interventions	17
Discussion	18
Constraints to girls' education	19
Potential solutions	19
Program costs.....	20
How do the findings of this review relate to other reviews?	21
Lessons for scaling.....	21
Conclusion	22
Appendix to "Girls' Education at Scale," by Evans, Mendez Acosta, and Yuan.....	24
References	29

Introduction

Gender equality is a stated objective in much of the world: indeed, the fifth Sustainable Development Goal is “achieve gender equality and empower all women and girls” (United Nations, 2015). Education is a crucial human capital investment that opens the door to subsequent economic opportunity. As a result, gender equality in education is one crucial step—albeit not the only one—towards achieving gender equality in life outcomes more broadly.

Girls’ education is often touted as one of the best investments in international development (Kim, 2016). But across low- and middle-income countries, adult women on average still have less education than men. Among young women and men in their early 20s, girls in South Asia and Sub-Saharan Africa still have less educational opportunity, whereas in other regions, girls have gained more ground (Evans et al., 2021a). These average shifts mask important differences across countries, within regions of countries, and across levels of schooling.

How to achieve gender equality in education at scale? Evidence on how to expand and improve girls’ education in low- and middle-income countries has expanded dramatically in recent years (Cameron et al., 2016; Sabet and Brown, 2018; World Bank, 2018). This review examines evidence from large-scale interventions, usually implemented by or in partnership with the government, to improve girls’ education. It focuses on studies demonstrated to improve either student learning or school completion, as opposed to more intermediate outcomes such as attendance or enrollment. It also discusses the quality of the evidence, where and how different solutions may apply differently, and signals where more evidence may be needed.

Our results show that programs and policies that have increased school completion or boosted learning for girls at scale in areas where girls face educational disadvantage include, among others, the elimination of fees or providing scholarships or stipends, reducing the distance to school or facilitating travel to school, providing school meals, improving the pedagogy of teachers through a range of inputs, and interventions that help students receive instruction at their level of learning.¹

We also discuss interventions that explicitly address issues faced principally by girls. These include sanitation and menstrual health, gender sensitization training, and safe spaces for girls. However, most of these interventions either have not been implemented at large scale or have not been evaluated with a focus on educational outcomes like improved learning and school completion. Nevertheless, we provide evidence on the outcomes they do shift (e.g., girls’ mental health and in some cases, their post-school transitions). Future, large-scale evaluations of such interventions will allow a better understanding of how well such programs can be implemented at scale and whether they shift educational outcomes.

¹ In this study, we do not distinguish strictly between programs and policies, as policies (such as the elimination of school fees) are virtually always accompanied by a program (such as providing grants to schools to compensate for lost fees).

In the discussion section of the paper, we propose three considerations—constraints, solutions, and costs—for policymakers and their partners as they apply evidence across different contexts. We also discuss lessons for scaling interventions effectively.

These findings complement those of other, recent reviews related to girls' education (e.g., Evans and Yuan 2021), the girls' education section of Evans and Mendez Acosta (2021), and the systematic review of interventions to improve girls' education by Psaki et al. (forthcoming). (Appendix Table A1 provides a summary of findings from five different reviews.) The current paper's focus on at-scale programs gives greater salience to programs that have been implemented by multiple countries nationwide, such as school fee elimination, school construction, or school meals.

Interventions to improve girls' education cannot ignore the ongoing global COVID-19 pandemic. The crisis has introduced large challenges for education: for example, as of early 2021, schools in South Sudan had been closed for 16 percent of a child's average lifetime schooling careers (Evans et al., 2021b). There are various channels by which the pandemic may be particularly harmful for girls' education: with a higher burden of housework while schools are closed, greater risks from possible adolescent pregnancies, and discriminatory treatment when resources for education are reduced (Kwauk et al., 2021; Mendez Acosta and Evans, 2020). Some of the available evidence indicates that drop-out in general may not be worse for girls than for boys, at least in Ethiopia (Kim et al., 2021), Ghana (Abreh et al., 2021) and Pakistan (Crawford et al., 2021). However, adolescent girls may be more vulnerable: a survey of almost 4,000 children age 10-19 years old in Kenya shows that girls were twice more likely not to return when schools reopened in January 2021 (Presidential Policy and Strategy Unit Kenya and Population Council, 2021). School fees were the most often cited reason, followed by unintended pregnancies (for girls). Learning loss has been severe for children in general, but even worse for girls in some countries. In South Africa, Grade 2 girls experienced a learning loss of about nine words per minute compared to losses of six words per minute for boys (Ardington et al., 2021). In Brazil, girls were less likely to take the end of the year standardized test in 2020 than the previous year, and were more likely to experience negative effects of remote learning in general (Lichand et al., 2021). Access and learning gaps exacerbated by the pandemic will require additional efforts to improve and expand girls' education at scale, and lessons from this paper can serve as a useful toolkit.

The challenges

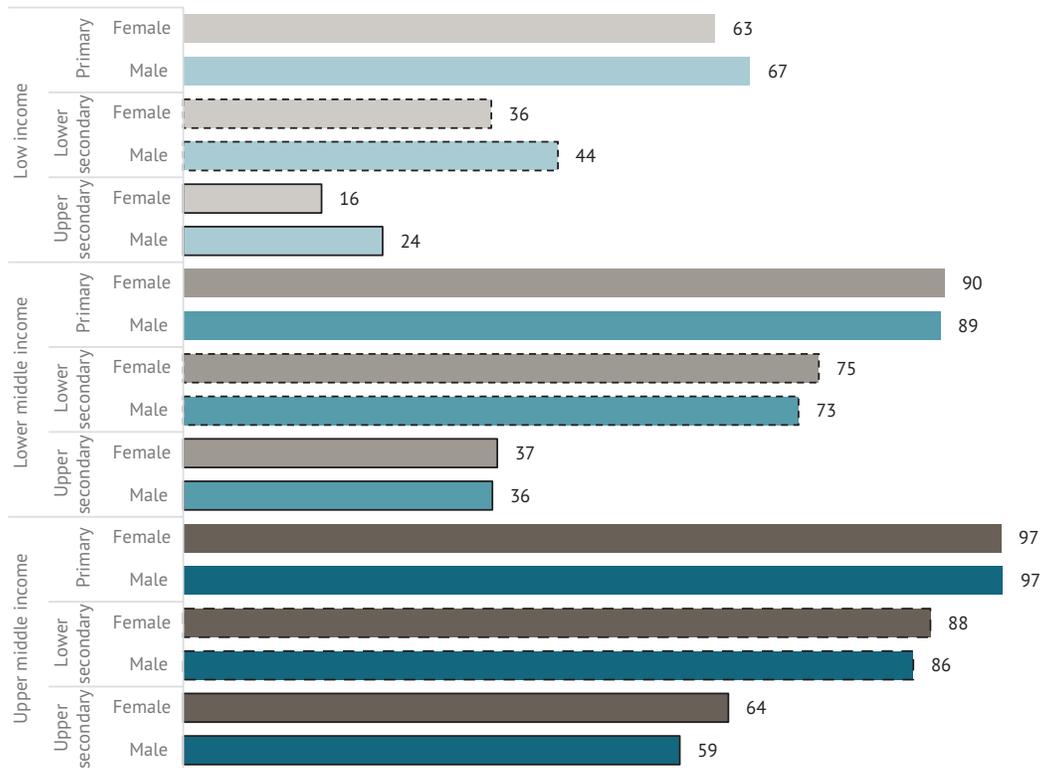
The broad policy challenge

Inequality in educational attainment is a massive global challenge, but the nature of the challenge varies dramatically across settings (Figure 1). For example, in low-income countries (like Afghanistan or Mali), boys are more likely to complete primary, lower secondary, and upper secondary education than girls. The gap grows with each level of education, doubling

from a four percentage point gap in primary school (63 percent versus 67 percent) to an eight percentage point gap in lower and upper secondary school.

In lower-middle income countries (like Bolivia or Ghana), girls and boys complete their basic education at essentially the same rates. But in upper-middle income countries (like Malaysia or Mexico), while there is parity in primary school completion (at 97 percent—almost every child is completing primary school), girls are five percentage points more likely to complete upper secondary school than boys. As countries grow more prosperous, gender gaps favoring boys disappear and girls even begin to pull ahead.

Figure 1. Gender inequality in education as measured by completion rates differ at the primary level versus the secondary level, and they vary in low-income versus middle-income countries



Source: Authors’ construction. The school completion rates for this figure are from the World Development Indicators (most recent available year) for primary and lower secondary education, and they are from UNESCO for upper secondary education. The data were downloaded in January 2021.

There are exceptions at every level of income. In Madagascar and Burkina Faso, both low-income countries, girls are more likely to complete primary and lower secondary school than boys. Angola, Benin, and Pakistan are lower-middle income countries with a remaining sizeable gender gap favoring boys at the primary level. In some upper-middle income countries such as Bulgaria, Gambia, and Guatemala, boys are at least three percentage points more likely to finish upper secondary schooling than girls.

Beyond access, there are also differences in learning outcomes. The World Bank’s harmonized test scores show that girls tend to have lower test scores than boys in low-income countries, with a concentration in Sub-Saharan Africa. In most middle-income countries, including most countries in Latin America, girls outperform boys on exams (World Bank, 2020b). Likewise, beyond differences in the gender gap across national levels of income, there will be differences in the gender gap across income levels within countries, between rural and urban areas (Evans, 2019), and across other vulnerabilities such as disability or orphanhood (Carvalho et al., 2021). This array of parameters across which gender inequality can linger explains why a middle-income country, despite having achieved gender parity on average, will still need to worry about gender equality in education.

At the same time, the specific challenges that girls (and boys) face change over time. For example, in a study of improved sanitation in Indian schools, Adukia (2017) found that the gains from sex-specific latrines principally appeared once girls had already reached puberty; for younger girls, the presence of a latrine helped, but whether or not it was sex-specific mattered less. Concern about sexual violence may be concentrated (albeit not limited to) older students. So just as countries will differ in whether they need to principally focus on closing gaps favoring boys or girls, the steps will vary based on the level of schooling.

This all reminds us that the challenge of gender equitable education is not a single challenge. Countries vary enormously in whether boys or girls are ahead in education and therefore need special attention and resources. Furthermore, even in countries where outcomes are similar, there may be important differences that require distinct attention. While this study explores—principally—interventions that have sought to improve girls’ education at large scale, no one intervention will apply in all cases.

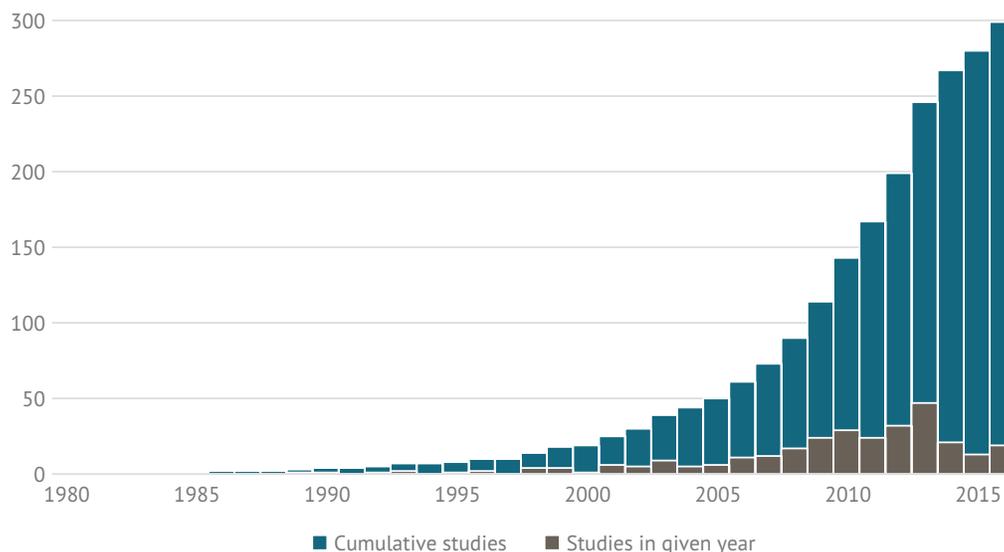
The knowledge challenge

The growth in evidence from evaluations of interventions on how to improve education has been dramatic in recent years, with a 15-fold increase in studies between 2000 and 2016 (Figure 2). Many of these evaluations either focus on girls education or present evidence on girls education within the context of a program that benefits both boys and girls (Evans and Yuan, 2021). Beyond methodological differences (e.g., experimental versus quasi-experimental evaluations), this evidence includes evaluations of various types of programs. Policymakers and partners can learn different things from each type of program evaluation.

One way to categorize the programs evaluated is based on implementer and scale, yielding four types: (1) pilot interventions implemented by non-government actors, (2) pilot interventions implemented by government actors, (3) large-scale interventions implemented by non-government actors, and (4) large-scale interventions implemented by government actors (Figure 3). For the sake of this paper, interventions that reach at least 10,000 students or were implemented nationwide will be considered “large-scale.” We recognize that in some countries, 10,000 is a significant proportion of students whereas in larger countries, that is not the case. While the number is ultimately an arbitrary cut-off, it does allow a distinction between truly small programs and those that require more management infrastructure and

resources to implement. Furthermore, the line between government and non-government implementation is sometimes blurry, with non-government support for implementation in government schools.

Figure 2. The growth in cumulative experimental and quasi-experimental education impact evaluations over time



Source: Authors' construction. Adapted from Figure S4.1 from World Bank (2018).

An example of the first type of study—pilot interventions implemented by non-government actors—might be the enrollment of about 2,000 children in alternative schools established by a non-government organization (NGO) in Guinea-Bissau (Fazzio et al., 2020). That intervention led to a six-fold increase in girls' test scores. In that context, government provision of education is very limited, and this study demonstrates that a holistic intervention—providing non-government schools with custom-designed teacher training materials and classroom materials for students and teachers, together with monitoring and evaluation of teachers and community outreach—can deliver dramatic gains in learning in an area with historically low learning levels. Another example of such a program evaluation examines the impact of setting up parent-teacher conferences in a study of about 4,000 students in Bangladesh (Islam, 2019). The program was implemented by a local NGO, and it more than doubled girls' test scores by the end of two years. Both studies provide promising interventions to increase girls' learning, but neither tells us whether it would be possible to implement such a program at scale. Evaluations of this style of intervention can be designed to inform scale up, but how well the program will actually work with thousands more students remains a challenge to know with certainty (Banerjee et al., 2017).

The second type of study implements an intervention at relatively small scale but using government systems. For example, the Government of South Africa implemented a randomized controlled trial (RCT) in 180 public primary schools, comparing the provision of traditional teacher professional development with the more interactive, on-site coaching. The coaching boosted girls' test scores four times as much as the traditional training

(Cilliers et al., 2019).² Because this was implemented through government channels, we can be more confident that it is possible to implement using teachers who have been recruited and are remunerated and managed through the government system. This points to a greater confidence that the intervention may scale (Gove et al., 2017).

Figure 3. What policy makers and partners learn from different types of program evaluations

		Scale	
		<i>Small</i>	<i>Large</i>
Implementation	Non-government	Does the theory behind the intervention apply in real life and in this setting?	Is it possible to reach large numbers of children and youth with the program?
	Government	Can existing structures implement this program under careful supervision?	Is it possible to reach large numbers of children and youth through existing structures?

Source: Authors' construction.

However, even within government systems, implementing a program on a large scale still poses new challenges (Anderson, 2018). First, the quality of implementers may suffer. Another government-implemented pilot impact evaluation in South Africa provided teachers with virtual coaching, i.e., a coach who reached out to teachers via tablet. Because the coach did not travel to schools, one coach was able to provide virtual support to many schools (Cilliers et al., 2021). But if such a program were scaled up nationally, more coaches would be needed. Would it be possible to find many coaches of similar quality of the pilot coach? In some places, maybe not. Second, the quality of supervision may suffer. Providing careful supervision to a pilot with a dozen or even a hundred schools is a different endeavor than providing careful supervision to thousands of schools. Third, in pilot programs it may be relatively easy to adjust the program as new challenges arise. Large-scale programs lose some of that quick flexibility. Fourth, programs often change models at scale to reduce costs or because of political pressures. A literacy program in Uganda was highly effective at boosting student learning when implemented by an NGO but was completely ineffective when a reduced cost version was implemented by civil servants (Kerwin and Thornton, 2021). A program putting teachers on short-term contracts led to learning gains when implemented by an NGO, but its design was compromised when implemented at scale due to political pressures (Bold et al., 2018; Duflo et al., 2015).

² The separated boy-girl impacts are not reported in the study, but they were communicated to us by the authors. For girls, the impact of coaching on literacy was 0.15 standard deviations ($p = 0.10$) and the impact of traditional training was 0.04 SDs.

Because of these and other challenges, large-scale programs often report smaller impacts than pilot programs, as has been demonstrated both in teacher coaching programs in high-income countries (Kraft and Blazar, 2018) and across a variety of education programs in low- and middle-income countries (Evans and Yuan, 2020).³ Thus, rigorous evaluation evidence of programs at scale is uniquely valuable.

Studies of the third or fourth type—implemented at scale, either in government systems or through NGOs that have the capacity to work at large scale—provide the most direct evidence of effective, at-scale interventions. Most of these evaluations are quasi-experimental. For example, an evaluation of a government program providing funds for bicycles to 160,000 secondary school girls in India (Bihar state) compares the girls who received bicycles to girls in a neighboring state and to local boys, who were not eligible for the program, a method called “triple differences” (Muralidharan and Prakash, 2017). The aforementioned study of latrine construction in Indian schools compares changes in outcomes among students attending schools where latrines were built through a large-scale government program (Adukia, 2017). These evaluations show that it is possible to implement the program at scale and still achieve significant impact.

While it is possible to learn from all these classes of evaluations, the focus of this study will be on studies that fall into the latter two categories, especially those that are implemented by government at scale.⁴ Most education is provided by the public sector: across low- and middle-income countries, less than 20 percent of primary education and less than 30 percent of secondary education is provided privately (World Bank, 2020c).⁵ As a result, interventions at large scale may be most sustainable if implemented through public sector mechanisms. NGOs that achieve results at large scale are also a key source of information for both policy makers and donors.

The methods used for this review

This is a narrative review of evidence of large-scale efforts to improve girls’ education at scale. The research team searched for evaluations that met two criteria: those that report impacts (either completion or learning) for girls and those that were implemented at large scale (at least 10,000 beneficiaries or nationwide implementation). The evaluations were drawn from various sources, including previous reviews of evidence on girls’ education that included both small and large-scale studies (Awosom et al., 2020; Evans and Yuan, 2021;

³ DellaVigna and Linos (2021) highlight that this pattern may be driven by “publication bias,” in which statistically significant results are more likely to be submitted and accepted for publication in academic journals (see section 4.3.1 of their study). Studies with smaller samples have less statistical power to estimate an effect, so smaller studies will on average require larger effects to show statistical significance.

⁴ Previous reviews have examined the full array of evidence. See Evans and Yuan (2021) for one example.

⁵ The numbers are 19 percent for primary and 28 percent for secondary. This is the low- and middle-income country aggregate provided for 2019, the most recent year for which data are available. This represents an increase from ten years previously, when the numbers were 15 percent and 23 percent for primary and secondary.

Sperling et al., 2016), the Millions Learning initiative (Robinson et al., 2019), the J-PAL post-primary education initiative, the “Learning @ Scale” initiative (Stern et al., 2020), and reviews on individual education topics (e.g., Read and Atinc 2018). The team included both evaluations that target girls specifically and evaluations that target both boys and girls but that report impacts separately for girls, or studies for which an earlier review identified gender-separated effects that were not reported in the original studies (Evans and Yuan, 2021).⁶

We reviewed and encoded the type of outcomes reported by the studies. The principal outcomes of interest were those nearer the end of the education production process: school completion (primary, secondary, and general educational attainment) and learning outcomes (literacy, numeracy, and related subject test scores). Other outcomes, such as enrollment or attendance, were included only insofar as they were instrumental to learning or to school completion, or in cases where other studies of similar programs established impacts on learning or completion. Subsequent life outcomes—e.g., income, fertility, or employment—were included in the relatively rare cases that they were available. We include studies with experimental designs (which are increasingly implemented at large scale but may be impossible to implement for national programs) and with credible quasi-experimental designs (such as difference-in-differences with matching). We also organized the studies by general categories of intervention. We reviewed this collection of studies and synthesized the findings.

We include a discussion of studies that impressively boost educational outcomes for girls in areas where girls remain disadvantaged in school—particularly Sub-Saharan Africa and South Asia (Evans et al., 2021a)—even in cases where both girls’ and boys’ education increase together. In those cases, similar absolute gains may reduce inequality if girls begin at a lower level: i.e., an increase in primary school completion of ten percentage points represents a higher percentage increase for girls if they begin at a lower level of completion. Even in cases where the percentage gains are similar, sizeable and significant gains to the quality and accessibility of girls’ education is likely to benefit girls regardless of impacts on boys. Finally, we discuss the evidence for interventions that commonly come up in discussions of girls’ education, such as sanitation and menstrual health product provision, in part to highlight which of those areas lack evidence at scale or with final educational outcomes like completion or learning.

We provide a narrative discussion of the effects. In this review, we do not standardize effect sizes across interventions: standard deviations in outcomes, while useful, can vary widely due to factors separate from the impact of the intervention such as the underlying distribution of initial value of the outcome in the study population, as well as—in the case of measuring learning—differences in the difficulty of changing the specific skills measured or different test designs (Evans and Yuan, 2020). Our approach of drawing studies from reviews (rather than

⁶ Previous research demonstrates that in some circumstances, general interventions deliver comparable gains to targeted interventions (Evans and Yuan) 2021. The choice of targeted or general interventions will depend on the principal obstacles girls face and whether those obstacles are principally felt by girls or not.

a time-consuming original search) and relying on narrative review rather than formal meta-analysis is consistent with documented rapid review methods (Tricco et al., 2017).

While we organize the interventions and programs as different groups of solutions, they could analogously be expressed according to the barriers they are addressing, namely (i) lack of financial resources to pay for direct and indirect costs of schooling, (ii) limited physical access to schools (iii) poor quality of education and limited academic support especially for girls falling behind, and (iv) barriers that are specific to girls such as limited access to sanitation hygiene and gender insensitive environment. There are more barriers to education that girls face than shown in this list, and the ones we present are those for which we have enough evidence available to draw insights (see Psaki et al. 2021 for a comprehensive catalogue of these barriers). We have also included a final section of results addressing some of the interventions with less or mixed evidence available.

The solutions

Make school cheaper

Fee elimination and scholarships

Many studies demonstrate that eliminating school fees or providing scholarships can dramatically increase school completion as well as learning outcomes. This applies in both primary and secondary school. While most countries have officially eliminated school fees at the primary level and some have eliminated school fees at the secondary level, in practice, families are often required to pay some sort of contributions to the school, above and beyond the cost of school materials and transportation and the opportunity cost of sending children and youth to school (Williams et al., 2015).

At the secondary school level, providing scholarships for youth in Ghana who had passed the secondary school entrance exam but who did not have the resources to pay—so keep in mind that this is a select group of students—led to more than a 60 percent increase in senior secondary school completion for girls (an increase of 25 percentage points relative to 40 percentage points in the comparison group). Girls were also much more likely to be enrolled in tertiary, although few were enrolled to begin with: that increased from 8 percent to 12 percent. The scholarships also led to a range of other positive impacts: better test scores in both reading and mathematics, better national political knowledge, media engagement, and a higher likelihood of having a bank account. Girls even had fewer pregnancies (Duflo et al., 2021). Eliminating school fees for secondary school girls in the Gambia increased the number of girls taking the high school exit exam (one proxy for completion) by 55 percent (Blimpo et al., 2019). Likewise, providing vouchers to cover the cost of private secondary school in Colombia—a program that reached 125,000 children—increased both test scores and secondary school completion rates, at comparable rates for both girls and boys (Angrist et al., 2006). Another scholarship program—for more than

100,000 girls in upper primary and lower secondary grades in the Democratic Republic of the Congo—boosted both reading and mathematics scores (Randall and Garcia, 2020).

Another program that paid school fees—this time in Tanzania—also covered other, informal costs for tens of thousands of secondary school girls who had been identified by their communities as highly vulnerable. The program also provided other benefits, such as textbooks and life skills training. Beneficiary girls had much higher test scores than girls at comparison schools. Less poor girls who attended beneficiary schools but did not receive financial support also had higher test scores, as did boys, suggesting a positive spillover effect. Girls who had their fees paid were 25 percent less likely to drop out of high school (Sabates et al., 2020). The evaluation of this program matched beneficiary girls with girls at other schools based on observable characteristics, so we are slightly less confident of the causal impact claim—i.e., it is possible that students in beneficiary schools were different than their comparators in ways we do not observe but that affect educational outcomes. Still, the results are consistent with evidence elsewhere that eliminating fees can boost learning and reduce dropout.

Feed children at school

There is a long history of evaluation evidence demonstrating that school meals boost enrollment at school. A large-scale school meals program in Pakistan—reaching hundreds of thousands of girls boosted school enrollment by forty percent (Pappas et al., 2008). There is some evidence that children across 32 African countries benefiting from a World Food Program school feeding initiative were enrolled in school at higher rates, with 27 percent higher gains for girls than for boys (Gelli et al., 2007).

A new generation of evidence demonstrates that school meals boost student learning as well. In Ghana, a large-scale school feeding program for which funding is now integrated into the government budget was evaluated via randomized controlled trial. After two years, both math and literacy scores rose for all children on average, but the largest impacts were for girls and for children in poverty. In other words, school feeding boosted learning especially for girls (Aurino et al., 2020). Likewise, a large-scale midday meal program in India led to improved test scores in both math and reading: in that case, girls and boys benefitted equally, as did poorer and less poor children (Chakraborty and Jayaraman, 2019). All of these interventions are implemented at the primary school level, although some smaller school feeding interventions do target adolescents (Drake et al., 2017). School meals also have well-documented health benefits, including reducing malnutrition, and it may make sense to see them more as social safety programs than purely as educational investments (Alderman and Bundy, 2012).

Cash transfers

Cash transfers are a widely used tool to achieve multiple ends: often the primary goal is that of a social safety net—ensuring that low-income households have money for essential needs—but they are often paired with further objectives related to health and education,

either explicitly through conditions that households must fulfill to receive the transfers or more subtly through labeling and encouragement (Benhassine et al., 2015). Many of these programs have been implemented at scale, and there are many variations, including those that combine a transfer with the direct payment of school fees, as in Bangladesh (Schurmann, 2009). There have been many evaluations of cash transfer programs on education, and most of those (nearly two-thirds) do report impacts separately for girls (Evans et al., 2020). However, one recent, fairly comprehensive review identified only a handful of studies that reported test scores or grade completion (Baird et al., 2014). That review reports consistent positive impacts on school enrollment—with higher impacts for conditional programs—but small impacts on test scores. Indeed, a recent report rates cash transfers as a “bad buy” if the objective is to boost learning (Global Education Evidence Advisory Panel, 2020). That said, ensuring that youth are in school is a likely precondition for learning and completion outcomes, so while cash transfers are not an effective instrument for learning alone, they may be crucial to making sure the most vulnerable children have the opportunity to reap learning gains from other interventions.

Make school more physically accessible

Two classes of interventions to make schools more accessible—constructing schools and providing transportation—have been implemented at scale with success and have shown positive impacts. Indeed, the Global Education Evidence Advisory Panel’s Smart Buys report, which draws on cost-effectiveness evidence, identifies reducing travel times to schools as one of its “good buys” (its second best rating, after “great buys”) for boosting learning in general, citing school construction and school transportation as two tested instruments (Global Education Evidence Advisory Panel, 2020). The benefits are experienced disproportionately by girls. Distance to school is a major constraint for many girls, especially at the secondary school level. For example, teenage girls in India who live 15 kilometers from a school have been more than one-third less likely to attend than those who live near a school (Muralidharan and Prakash, 2017).

One solution to that challenge is to build schools close to girls. Several interventions that have employed this approach have constructed schools with the needs of girls in mind (i.e., “girl-friendly” schools). A program that constructed primary schools in Burkina Faso benefited many thousands of children (the precise number is unclear): after 2.5 years, enrollment and test scores rose, and both of these effects were greater for girls than for boys (Kazianga et al., 2013). After ten years, impacts on test scores and enrollment remained although they were smaller, which may be unsurprising because many comparison villages also had some sort of school. However, primary completion rates for girls were more than double in beneficiary villages than in comparison villages (23 percent versus 9 percent); they were also much higher for boys (39 percent versus 30 percent). Marriage rates for girls were also much lower in beneficiary villages (33 percent versus 39 percent). The schools all had separate latrines for boys and girls and a water source. Importantly, the program did not merely construct schools: it also provided school meals, books, and an information campaign to parents on the importance of education (Davis et al., 2016). A slightly smaller scale program in Niger boosted student test scores, but only for girls (Bagby et al., 2016).

We know something about the long-run impacts of school construction. In Indonesia, the government implemented a massive school construction program: more than 61,000 schools were constructed between 1973 and 1979. Women were more likely to complete primary school (by 4 percentage points) as a result, and their children were more likely to complete secondary school and even tertiary education, with slightly larger effects on their daughters than on their sons (Akresh et al., 2019).

A second intervention extends existing schools. In several countries of Latin America (Brazil, Colombia, Ecuador, Honduras, and Nicaragua), a program provided alternative lower- and upper-secondary education to youth who otherwise would not have access. The program is “alternative” in the sense that it seeks to integrate academic learning with practice livelihood experience, and it has reached some 300,000 students (Kwauk and Robinson, 2016). An evaluation of the program in Honduras demonstrated higher test scores at lower cost than traditional schools (McEwan et al., 2015).

A third intervention involves making transportation to school more accessible. In India, a program provided cash transfers intended for the purchase of bicycles to more than 150,000 girls. School principals then provided receipts demonstrating that the cash had been used to purchase bicycles. The impacts of the program are striking: the gender gap in secondary school fell by 40 percent, and the girls who passed the high-stakes secondary school exam rose by 12 percent (Muralidharan and Prakash, 2017). A much smaller version of the program (benefiting several thousand girls) distributed bicycles to younger schoolgirls (in upper primary school); the evaluation found the program reduced absenteeism, commute time, and teasing but had less dramatic effects otherwise (Fiala et al., 2020).

Teach better

General improvements in the quality of instruction

In Kenya, a multi-faceted literacy and numeracy program was implemented through government systems. It included detailed teachers’ guides, training for teachers and head teachers, teacher coaching, and literacy and math books for every student. The pilot program—evaluated via a randomized controlled trial—led to impressive literacy gains in the early years of primary school, with comparable impacts for girls and boys. Impacts for mathematics were more modest but still positive (Piper and Mugenda, 2014). When the program was scaled up nationwide (an initiative called Tusome), literacy gains were still positive and sizeable, even though some aspects of implementation were not as well done as they were at the pilot. For example, teachers received some coaching in the scaled up version but less than planned (Piper et al., 2018). Gains were slightly bigger for girls than for boys on several of the literacy measures (Fraudenberger and Davis, 2017). A related program, implemented at large scale in Pakistan and evaluated using a quasi-experimental program, also provided detailed teachers’ guides and reading materials for students, among other activities (Chemonics International Inc., 2019). Beneficiary children boosted their reading more than non-beneficiary children, with larger gains for girls (Management Systems International, 2018). An early-grade literacy program that trained teachers and librarians, assigned literacy

coaches to work with teachers, and provided lesson plans and reading materials led to increased reading fluency gains for students in Laos and Zambia compared to students from comparison schools (Alexander et al., 2016). Another quasi-experimental evaluation of the same program, this time in India, showed similarly significant gains in reading fluency for both boys and girls (Joddar, 2018).

Target children who fall behind

In India, a remedial reading program implemented by Pratham, a large NGO, reaches hundreds of thousands of students. Hiring young women to teach students who are falling behind in their basic numeracy and literacy skills led to significant gains for both girls and boys (Banerjee et al., 2007). Pratham has implemented several variations on programs to bring learners who are falling behind up to speed, including intensive summer reading camps or using one hour of the school day to group learners by ability rather than official grade level (Banerjee et al., 2017).

In Ghana, the government implemented a related program with more than 80,000 students without NGO-support (beyond a visit to Pratham in India to see their programs in action). Schools hired teaching assistants, many of whom had no previous experience with teaching. In one model, the assistants pulled remedial learners out of class for part of the school day for special attention. In a second model, the assistants provided the same attention but after school hours. In a third, assistants worked with half of a class for part of the school day, effectively just reducing class size and allowing more specialized attention per student. In a fourth model, there were no teaching assistants, but teachers were trained to focus their teaching at whatever learning level students were at, rather than being bound by the curriculum for their grade. All four models led to significant gains in student learning, and three of the four (all except the third) were equally cost-effective. Gains were higher for girls than for boys (Duflo et al., 2020).

What about other teacher policies?

Teacher incentives in India, implemented in schools covering more than 20,000 students, led to sizeable, significant gains in language and especially math scores, with no significant differences for girls and boys (Muralidharan and Sundararaman, 2011). In another teacher performance pay intervention in Tanzania that reached more than 120,000 students, students in schools that received a combination of school grants and teacher performance incentives saw the largest test score gains, and the gains for girls were significantly larger than those for boys (Mbiti et al., 2019).⁷

⁷ Another class of teacher incentives provides additional financing to teachers who work in rural or otherwise disadvantaged schools. While those incentive programs, implemented at scale, tend to be effective in leading teachers to relocate, they did not boost girls' test scores in Zambia (Chelwa et al., 2019), although they did, particularly for higher income girls, in the Gambia (Pugatch and Schroeder, 2018).

A large-scale teacher professional development in China had no impact on student learning for girls (or for boys), arguably because the training was too theoretical (Loyalka et al., 2019). More broadly, at-scale teacher professional development programs tend not to incorporate the elements that are associated with the best student learning outcomes in smaller programs (Popova et al., 2021). An educational program in Tanzania with a major focus on teacher professional development, evaluated with a quasi-experimental design, led to significant gains in literacy and numeracy in the early grades of primary school, with the biggest gains for girls. While attribution across components is difficult, the authors hypothesize that the teacher professional development played a major role, as it was implemented effectively and led to increase use of teaching aids and boosted teacher confidence and motivation (Ruddle and Rawle, 2020).

Deploy effective forms of education technology to improve instruction

Many education systems invest extensively in education technology (or ed-tech), although reviews have found decidedly mixed impacts of education technology investments on student learning outcomes (Bulman and Fairlie, 2016; Escueta et al., 2020; Evans and Mendez Acosta, 2021; Rodriguez-Segura, 2021).⁸ The heterogeneity of impacts likely derives from the fact that technology plays many roles in education: technology can be used to improve the quality of instruction, to nudge parents or students, or for self-led learning. Some interventions to improve the quality of instruction have been successful. Many of these programs have been implemented at large scale, although the evaluations often study a smaller sample of students. For example, providing children with technology-aided after school instruction in urban India led to large gains in math and language scores that were comparably sized for boys and girls. The evaluation sample was fewer than one thousand students, but the software has been used by more than hundreds of thousands of students (Muralidharan et al., 2019). However, this presents a conundrum: a small evaluation may have much more controlled conditions than large-scale use of a software. In this case, a larger, in-school version of the program was implemented for one period per day and still delivered significant positive learning impacts, albeit smaller than those in the pilot, after school program (Muralidharan and Singh, 2021). Likewise, a large-scale computer-assisted learning platform in Uruguay that has reached at least half of all students in third through sixth grade delivered significant learning gains of comparable size for both girls and boys (Perera and Aboal, 2019). Both at-scale computer-assisted learning programs delivered similar results: roughly 0.2 standard deviations of mathematics learning, which amounts to a significant gain relative to a year's learning (Evans and Yuan, 2019).

There is less evidence on educational television, although there are several ongoing studies of educational television and radio in the context of the COVID crisis (World Bank, 2020a). The existing evidence suggests that high-quality educational television programming can boost cognitive outcomes for girls in particular, although that evidence is from younger

⁸ Some high-performing education systems in high-income countries have actually reduced investments in education technology relative to education systems in other high-income countries (Ripley, 2014).

children (Mares and Pan, 2013). That said, across African countries, even with schools closed during the pandemic, relatively few households have access to most ed-tech products (Crawford, 2020).

It is with some reserve that we include ed-tech among the solutions, since simply providing education technology does not inherently promise any gains. Indeed, several large-scale programs have provided computing equipment and had no impact on student learning or other outcomes. This result is important because scaling hardware programs, while costly, may appear logistically straightforward. But providing computers to more than 6,000 schools in Colombia had no impact on student learning for girls or for boys (Linden and Barrera-Osorio, 2009). Providing laptops (more than 1.6 million of them!) to children in Uruguay had no impact on girls' learning outcomes in either the short run or the longer run (Yanguas, 2020); a similar program that delivered tens of thousands of laptops in Peru had no impact on girls' learning nor on students' dropout rates (Cristia et al., 2014, 2017). Providing laptops or desktops may be a necessary step in providing technology-assisted learning, but programs that do the first without the second are poised for failure.

Many other technological innovations have been implemented only at smaller scale, with far fewer than 10,000 students (e.g., Berlinski et al. 2016 or Duflo et al. 2012). In other cases, ed-tech interventions have been implemented at large scale but lack serious evaluation on student outcomes, e.g., several large-scale education technology platforms implemented in India (Bajpai et al., 2019; Dhar et al., 2016).

Gender focused interventions

While many obstacles affect girls and boys differentially, several classes of interventions seek explicitly to address obstacles felt principally by girls. Many of these interventions are not focused on academic outcomes, and so do not report school completion or learning outcomes, although there are exceptions. But a lack of evidence on those outcomes does not mean these are not important investments. It may mean that the primary motivations for investing in some of these programs are not to boost learning outcomes or school completion but rather to protect girls, to boost their overall well-being, or to prepare them for life after school.

Sanitation and menstrual health products

Toilets for girls often arise in discussions of gender-equitable education. Indeed, a national school latrine construction program in India increased enrollment for girls. For the youngest girls, any latrine boosted enrollment. For girls who had reached puberty (i.e., upper primary), only sex-specific enrollment boosted outcomes. While the study did not measure results on completion, the enrollment results are strong and enduring three years after construction. Students did not perform better on direct tests of their learning, but girls (and boys as well) both sat for and passed their official school exams at higher rates (Adukia, 2017).

Two reasons that sex-specific latrines often come up in discussions is because of the fear that girls will either miss school during menstruation or that they might experience verbal or physical harassment. On the former point, estimates of menstruation-related absenteeism vary dramatically across contexts, with large estimated impacts in Bangladesh and India but low estimated impacts in Kenya, Malawi, Nepal, and Uganda (Benshaul-Tolonen et al., 2020b). In Kenya, providing sanitary pads (which are girls' choice) did reduce absenteeism, although those results did not hold up if one excludes girls who had transferred away from the study schools (Benshaul-Tolonen et al., 2021). In both Kenya and Nepal, providing an alternative, less familiar technology—a menstrual cup—did not affect absenteeism. (The cup is much cheaper, but an intervention cannot be cost-effective if it is not effective!) Regardless, these are both relatively small studies, with a few hundred students in Nepal and a few thousand in Kenya. However, looking beyond absenteeism, providing either pads or a cup in Kenya significantly improved emotional and social well-being (Benshaul-Tolonen et al., 2021); and in Tanzania, having insufficient menstrual materials was associated with more teasing of girls about their periods (Benshaul-Tolonen et al., 2020a). On net, the evidence for providing menstrual hygiene materials on a purely instrumental basis (getting girls to school completion or higher test scores) is still weak, although the emotional well-being of female students is another important, worthy end in itself.

Gender-sensitization training

When discussing gender equality in education, a class of programs that often arises in discussion is some sort of training to increase the sensitivity of teachers, school managers, or students to gender issues. To date, there is limited evidence evaluating such programs at large scale or on outcomes such as school completion or student learning. For example, one program evaluated the impact of 2.5 years of classroom discussions on the topic of gender equality among several thousand sixth and seventh graders in India. (The evaluation sample was about 14,000, with just under half of those students receiving the program.) While that is not a small-scale program, it also is not at a massive scale, and the outcomes measured are reported views on gender norms and some self-reported behaviors (e.g., boys reported a higher likelihood of doing chores). The program tested neither student learning (beyond on gender attitudes) nor school completion (Dhar et al., 2020).

Another program that functionally acted as a gender sensitization program was the Power to Lead program, which provided training leadership skills for girls across six countries. In practice, more than 30 percent of participants were boys, and these reported improved understanding of gendered social norms (Baric, 2013). Among girls, a mostly qualitative evaluation found positive impacts on measures of leadership skills, leadership action, and self-confidence (Miske Witt & Associates Inc., 2011).

The fact that none of these studies measured school completion or student academic learning is not a critique: those were not the principal objectives of the programs, and reducing sexism is a valuable objective in its own right. But it may also have instrumental value, increasing completion or student learning—for example, if girls are able to focus more on their studies in a more gender sensitive environment (e.g., with less teasing or harassment

or with a more gender-equal distribution of household tasks). Further research will be required to see if this is in fact the case.

Mentoring and safe spaces for girls

Creating places where female students can engage without boys or men is sometimes proposed as a useful intervention (Megevand and Marchesini, 2020). Indeed, evidence on this type of intervention is promising, although the outcomes measure are not school completion or student test scores. For example, a program that formed clubs for more than 50,000 adolescent girls in Uganda and provided vocational training and information about sex, reproduction, and marriage, led to reduced adolescent pregnancy and more engagement in income-generating activities four years later (Bandiera et al., 2020a). However, an adaptation of the same program in Tanzania had no impact on similar outcomes (Buehren et al., 2017).

A program forming girls' clubs in Sierra Leone—with a smaller sample of 150 beneficiary villages—was interrupted by school closures due to the 2014/2015 Ebola epidemic; but five years later, girls in communities where girls' clubs had initially been formed were much less likely to have experienced an adolescent pregnancy and much more likely to have re-enrolled in school after the epidemic (Bandiera et al., 2020b). These are important outcomes beyond education; the impact of these types of programs on purely educational outcomes are less well known at a large scale. A small scale, government implemented program in 20 low-performing secondary schools in Trinidad and Tobago converted co-education schools to single sex schools: girls subsequently performed better on secondary school completion exams, and both boys and girls took more advanced coursework (Jackson, 2019).

An empowerment intervention assigned “social mobilizers” to schools to provide life skills classes and mentoring and has reached over 95,000 adolescent girls in nine countries. A clustered randomized evaluation of the program in India showed positive effects on drop-out rates and life skills such as future planning, empowerment, and attitude towards gender norms, although those gains did not translate to child labor outcomes or test scores (Edmonds et al., 2019).

Other interventions

Effective interventions with less evidence behind them

Some interventions have proven effective at scale but have been rigorously evaluated in only one or two settings, so policymakers and partners may feel less confident that their success can be replicated elsewhere. For example, a program that provided eyeglasses to almost 30,000 students in China boosted test scores equivalent to nearly a full year of business-as-usual schooling for students with poor vision, benefiting both girls and boys, although one-third of girls refused eyeglasses while only one-fourth of boys did (Glewwe et al., 2016). Likewise, a program in India—implemented by the NGO Pratham (like some of the literacy programs above)—reached a large sample of mothers to either provide literacy for mothers, train mothers in how to boost their children's literacy, or both. All three variations

of the program had positive (but modest) impacts on children’s mathematical performance (Banerji et al., 2017). Of course, providing eyeglasses to children who need them is a worthy and important objective, as is making sure that mothers can support their children at home. But the question is not the worthiness of the objective but rather how well such programs can be implemented at scale and the gains relative to other programs in terms of outcomes of interest.

Areas with more limited, mixed evidence

School accountability, often involving publicizing either resources flowing to schools or student performance in schools, is an important area. In Uganda, using newspapers to publicize the amount of funds that schools would be receiving from the central government boosted student learning outcomes, with apparently larger effects for girls (Reinikka and Svensson, 2011). But programs publicizing results—at scale—have been less effective. For example, a program in India that facilitated meetings to discuss education and—in some cases—invited community members to create “report cards” on learning status in the community had no impact on learning outcomes for girls (or boys) (Banerjee et al., 2010). In Tanzania, a nationwide, low-stakes accountability program published school rankings: while it boosted learning outcomes in the poorest performing schools, it also led those schools to exclude students—both girls and boys, to equal degrees—from their last year of schooling, an unfortunate, unintended consequence (Cilliers et al., 2020). Many school accountability programs that reached large numbers of students do not report impacts separately for girls (e.g., Barr et al. 2012 in Uganda, Pandey et al. 2008 in India, Andrabi et al. 2017 in Pakistan, and many others).

Discussion

In this review, we have presented various initiatives that have been implemented at large scale, usually through government channels, resulting in large learning or completion gains, especially for girls. However, the reader may ask: which of these initiatives is the best bet? Unfortunately for anyone seeking a silver bullet, the answer comes down to the economist’s favorite answer: it depends. Let us demonstrate with two—hopefully obvious—examples.

When is school construction an effective intervention? In Benin in the 1990s, the government or NGOs constructed more than 1,500 new schools, and enrollment surged by almost 200,000, driven mostly by girls. But the increase in enrollment was concentrated in rural areas, where there were fewer schools before the construction boom (Deschênes and Hotte, 2019). The interventions where school construction was effective, in Burkina Faso and Niger and Indonesia, were in locations or times when schools were scarcer. Is school construction a good bet? Yes, if there are few schools.

When is school feeding an effective intervention? In contrast to the examples we highlighted above, a large-scale school feeding program in Chile had no impact on learning outcomes or grade progression (McEwan, 2013). A program in Sri Lanka similarly showed no impacts.

Why not? Chile had already eliminated extreme malnutrition and educational outcomes were relatively strong. Likewise, Sri Lanka already had high rates of enrollment (Snilstveit et al., 2016). School feeding is a powerful tool, but only in places where children face this particular need. Is school feeding a good bet? Yes, if children are malnourished. Ultimately, the most effective intervention in a given location will depend on the circumstances of that location.

Thus, how should an education system decide which interventions to invest in to boost girls' education at scale? We propose three aspects for consideration: constraints to girls' education, potential solutions, and program costs.

Constraints to girls' education

First, education systems need to identify their key constraints. What are the weaknesses in the education system? What is holding girls' education back? In recent years, a range of diagnostic tools have been deployed that have demonstrated challenges in the education system overall, and some can be deployed towards understanding girls' education. Several tools measure student learning, including citizen-led assessments like ASER and Uwezo, along with school-based assessments included in the Service Delivery Indicators and in national and regional exams (ASER, 2021; Uwezo, 2021; World Bank, 2021). These can help identify which regions face the biggest gaps in learning. Likewise, high quality gender-disaggregated systems data can track student school completion rates. In terms of understanding the reasons behind high student dropout or poor student learning, some household surveys (e.g., some of the Demographic and Health Surveys or the Living Standards Measurement Surveys) ask directly about reasons that girls drop out of school. Furthermore, the Service Delivery Indicators measure the health of the teacher workforce, with a focus on skills and absenteeism. If a survey demonstrates that teacher absenteeism (or any other issue) is an important problem, additional diagnostics may be needed to understand the reasons behind the issue, in order to design the most effective mechanisms.

Other diagnostic tools seek to measure the quality of the overall education system. They may be helpful in reviewing the inputs, resources, different actors and institutions, and the politics and feedback loops to improve education systems for girls (Faul, 2016; Savage and Martinez, 2019). A review of country-level efforts to promote policies that are more supportive for girls education, such as the Girl's Education Policy Index (Crawford and Hares, 2020), can help to identify systemic exclusionary policies. Public expenditure tracking surveys measure how well resources reach schools and can inform education priorities.

Potential solutions

With the rapid increase of evaluation evidence, education systems have many solutions available to them. What evidence should policymakers draw on in selecting the best investments to boost girls' education at scale? If there is high-quality local evidence on

effective solutions, then that can be an excellent source.⁹ If not, then Bates and Glennerster (2017) propose a four-step framework for deciding if evidence of an effective program in one place will apply in another: (1) understand the theory behind the original program; (2) verify that conditions in the new location hold for the same theory to apply; (3) weigh the strength of the evidence that the same mechanism would work to change behavior in the new location; and (4) determine the likelihood that the program can be effectively implemented in the new location. This process involves drawing on a mix of the most rigorous evidence from anywhere and the best available local evidence.

No policy or program operates in a vacuum, so a key, iterative interplay between constraints and solutions will entail examining proposed solutions in the context of existing policies and how they are likely to interact. Effective overall education system reforms that deliver significant gains to girls as well as boys, like those documented in Finland over several decades (Sahlberg et al., 2021) or the Brazilian state of Ceará over a shorter period (Loureiro et al., 2020), require a collection of solutions.

Like all reviews, this study is limited by those areas that have been evaluated. Education systems should continue to innovate. Some of that innovation may be in adapting within areas already shown to be effective. Reducing education costs for girls, for example, may be accomplished in various ways. Other innovation may be completely new. However, there are areas that have not been effective at scale across multiple settings, like distribution of computer equipment either without plans or capability to integrate it fully into the system or without accompanying investments in the complementary technologies needed to deliver gains in learning. These should be avoided.

Program costs

Every new program and most new policies come with a price tag. Ultimately, we care about both effectiveness in delivering gender equality in education and about cost-effectiveness. By definition, the most cost-effective interventions deliver the biggest gains per dollar spent. But cost-effective interventions that deliver small gains, while often worth doing because of low costs, will not be sufficient to close gender gaps. So an information campaign that costs little may be worth doing because of a high benefit-cost ratio. A school feeding program or a school fee elimination program will cost much more but may—depending on the constraints—deliver larger gains as well. (Programs like cash transfers and school feeding programs may appear less cost effective purely in terms of education gains because many of their benefits extend beyond the education sector.)

⁹ If there is local evidence, Pritchett and Sandefur (2013, 2015) argue that less rigorous local evidence may be more relevant than more rigorous evidence from another context. This requires at least two caveats, however. First, in many cases, there is no relevant local evidence of impact (even non-rigorous evidence). Second, whether or not this is true will depend fundamentally on the size of the selection bias (which RCTs overcome) relative to the size of the impact of the program or policy. If the selection bias is large (e.g., poor people are more likely to get cash transfers and sickly people are more likely to go to the doctor), then non-experimental methods may be less informative.

A minority of evaluations report costs, but the proportion appears to be growing over time (Evans and Mendez Acosta, 2021; McEwan, 2015). Yet just like impact estimates, costs for the same program can vary significantly across contexts (Evans and Popova, 2016). Just as with the four-step approach for adapting impact estimates across contexts, education teams and those who support them will need to adapt cost estimates across contexts as well.

How do the findings of this review relate to other reviews?

Several other reviews have explored how to ensure education for girls, including Psaki et al. (forthcoming), Evans and Yuan (2021), Sperling et al. (2016), Unterhalter et al. (2014), and Tembon and Fort (2008). None of those reviews has focused on more final educational outcomes nor exclusively on programs at scale. In Appendix Table A1, we summarize the recommendations of each of those reviews. All previous reviews highlight the value of eliminating fees and of cash transfers, at least for access outcomes. Provision of food is rated as effective by Psaki et al. (forthcoming)—particularly for enrollment and attainment outcomes—and promising by Unterhalter et al. (2014), albeit with a caveat warning of potential adverse learning effects if feeding leads to overcrowding. As we highlight, subsequent research has demonstrated a consistent positive impact of feeding programs on learning. The majority of the reviews also highlight the role of teaching quality (i.e., pedagogy), especially efforts to provide academic support for those lagging behind (i.e., remedial education). Other reviews are more conclusive than ours on the role of sanitation, likely because we restrict to school completion, whereas most impacts measured for sanitation have been enrollment or absenteeism.

Psaki et al. (forthcoming) also highlight several areas where there are evidence gaps, including programs to reduce adolescent marriage, school-based health programs, and many others. The gaps are even more abundant when we limit ourselves to studies at scale, which is a reminder that there is still much room to innovate and evaluate. Some areas that seem like they could be effective lack any evidence at all: for example, information campaigns about the returns to education have been cost-effective at boosting test scores and increasing access (Angrist et al., 2020), but we are not aware of any evaluations of campaigns focused on the returns to girls' education, despite the fact that those might be a natural extension of the existing success with information dissemination. Psaki et al. (forthcoming) similarly highlight “efforts to increase support for girls' education” among commonly used approaches with little evidence.

Lessons for scaling

Earlier in the paper (The Knowledge Challenge), we discussed some of the challenges in taking interventions to scale. First, as programs expand, they often outgrow their monitoring, management, and evaluation systems, especially if the scale up involves new partners or a larger (maybe even different) government agency. In some instances, scaling up might mean creating new institutions or legal entities to manage the expansion effectively (Cull and McKenzie, 2020). Expansion also often requires revamped incentives and accountability

structures which in turn requires stepping up monitoring, evaluation, and oversight (Hartmann and Linn, 2008).

Second, political momentum and political champions are critical resources to ensure that the program both acquires and retains legitimacy to survive changes in administrations. In Ceará, Brazil, arguably the single most important factor in a successful education reform that dramatically boosted student learning outcomes was consistent political leadership (Loureiro et al., 2020). In addition, scaled up programs or national reforms also capture constituents' attention more than small pilots. In the case of interventions targeted exclusively to girls, this may generate backlash due to a perception of “female bias” (Subrahmanian, 2005). Political champions are key to providing platforms necessary for communicating to local leaders and ensuring compliance or acceptance.

Finally, special measures that deviate from existing schooling provision systems—such as alternative schools, early child education systems, or health and hygiene programs—are more likely to be sustained if they are integrated within the formal system of provision rather than standing alone as isolated efforts with separate management systems (Subrahmanian, 2005). Similarly, mixed packages of reforms which include simultaneous and complementary programs will increase the likelihood that the effect of one large-scale program can build on the opportunities created by other, similar reforms.

Conclusion

What this review has sought to do is highlight interventions for which there is evidence from multiple settings that they can be implemented effectively at large scale and deliver positive impacts for girls. These interventions can increase gender equality where girls are disadvantaged. Gender equality in education and gender equity in education are different, related concepts. Equality may be associated with achieving similar education outcomes for boys and for girls. Equity, rather, is associated with “fairness or justice in the provision of education” (Espinoza, 2007). Achieving gender equal outcomes in education (e.g, gender parity in school completion or learning) may require gender inequality in resources spent. One might also argue that, if women face greater challenges than men later in life, gender inequality in education may be needed to achieve gender equality in later life outcomes.

This review takes an “effects of causes” approach, where one starts with interventions or policies and exams the effects of those policies. An alternative approach would be to take a “causes of effects” approach, examining countries that have made great strides in girls' education and seeking to discern the causes thereof (Goertz and Mahoney, 2012). This approach has been applied to understanding the quality of education in high-performing countries, e.g, in Vietnam (London, 2021). Future work can explore the causes of large gains in girls' education, as some countries have made impressive strides in short periods of time (Evans et al., 2021a).

In this paper, we have also provided a discussion of gender-specific interventions and guidance as to how to make sense of the large and growing body of evidence. There is no guarantee that a given impact will replicate in every setting, but this collection of evidence provides a menu for policymakers and partners that comes one step closer to feasible implementation than previous reviews that draw on a higher proportion of small-scale, NGO-implemented interventions. Achieving gender equitable education is an ongoing challenge, but there are proven solutions that work at scale.

Appendix to “Girls’ Education at Scale,” by Evans, Mendez Acosta, and Yuan

Appendix Table A1. Evidence map based on reviews of interventions that focus on improving girls’ education

Reviews	Psaki et al. (forthcoming)	Evans and Yuan (2021)	Sperling et al. (2016)	Unterhalter et al. (2014)	Tembon and Fort (2008)
Coverage	“82 experimental and quasi-experimental studies (88 papers) that employ interventions or analyze the effects of exposures that address at least one gender-related barrier to schooling and measure impact on girls’ education outcomes” from 2004 to 2020	“270 educational interventions from 177 studies in 54 low- and middle-income countries” (both general and girl-targeted interventions) that report impacts on girls. Search covers studies published before 2018	138 studies that constitute “a very robust set of evidence of what works in girls’ education”	177 studies that are “direct or indirect intervention which address aspects of girls’ education and/or poverty and have been published since 1991.” They classify the evidence available for different interventions as strong, promising, limited, or more research needed.	Presents select interventions that are “proven successful in raising female enrollment and completion rates.”

How does each review rate the effectiveness of the following interventions?

ability to afford tuition and fees	Effective	Effective in improving girls’ access (cash transfer and subsidy)	Reducing tuition fees are effective in increasing girls’ enrollment, but gains are not enough for girls from especially poor households. Providing cash transfers and stipends are effective in reducing the indirect and opportunity costs of schooling.	Strong effect on school participation and performance (cash transfers)	Eliminating user fees and providing stipend and conditional cash transfer to girls are proven successful.
adequate food	Effective for enrollment/attainment outcomes			Promising effect of school feeding on enrollment but potentially negative effects on learning in the case of overcrowding	

Reviews	Psaki et al. (forthcoming)	Evans and Yuan (2021)	Sperling et al. (2016)	Unterhalter et al. (2014)	Tembon and Fort (2008)
access to school	Promising	Effective in improving girls' access (reducing commuting time by building schools closer to girls)	Building schools closer to girls dramatically improve attendance and test scores.	Provision of additional schools has strong effect on girls' enrollment and promising effect on learning outcomes More research needed on impact of school choice and inclusive strategies (availability of private schools might widen gender gap as boys get sent to better-resourced schools, differences in outcomes across single-sex vs co-educational and faith-based schools).	
adequate school materials	Promising				
water and sanitation in schools, especially toilets	Promising	Effective in improving girls' access (hygiene promotion and water treatment)	WASH programs are effective in reducing dropouts and school absences.	Promising impact of integrated water, sanitation and hygiene interventions	
school-related gender-based violence	Not enough studies (evidence gap)		Some promising evidence on increasing awareness and shifting norms are available. Girls' clubs and safe spaces can also improve gender-based violence outcomes.	More research needed on the impact of interventions aimed at reducing gender-based violence on school participation.	
sports programs for girls	Not enough studies (evidence gap)				

Reviews	Psaki et al. (forthcoming)	Evans and Yuan (2021)	Sperling et al. (2016)	Unterhalter et al. (2014)	Tembon and Fort (2008)
health and childcare services (school-based)	Not enough studies (evidence gap)	Effective in improving girls' access (malaria prevention)	Nutrition programs and deworming are effective in improving enrollment and attendance.	Promising effect of iodine supplements and deworming on enrollment but potentially negative effects on learning in the case of overcrowding Promising effect of teaching personal, health and social issues linked with sex education.	
child marriage and adolescent pregnancy	Not enough studies (evidence gap)				
menstrual hygiene management	Not enough studies (evidence gap)		Interventions that provided menstrual supplies had mixed evidence of impact on attendance or scores.	Menstrual hygiene interventions offer limited evidence of direct impact on girls' attendance	
life skills	Not enough studies (evidence gap): Most studies were multicomponent, and approaches to life skills varied.			Promising effect of women's literacy programs that also offer life skills and safe spaces.	
information on returns to education/ alternative roles for women	Not enough studies (evidence gap)			Providing information on returns to schooling has a strong effect on school participation	Raising education standards and quality to improve economic returns to female education is included in the list of successful interventions.
safe spaces and social connections	Not enough studies (evidence gap)		Child-friendly spaces and back-to-school programs are crucial in supporting girls' education during emergencies and conflict settings.	Girl-friendly schools has a strong effect on school participation.	

Reviews	Psaki et al. (forthcoming)	Evans and Yuan (2021)	Sperling et al. (2016)	Unterhalter et al. (2014)	Tembon and Fort (2008)
policy/legal environment	Not enough studies (evidence gap): Policies are varied and findings inconsistent.			More evidence needed on the effect of engaging networks of women activists in influencing gender-equal policy development More research needed to assess impact of infrastructure interventions combined with policy and institutional culture change towards girls' rights.	Focusing "attention on gender inequality by means of advocacy and better impact evaluation research" is proven successful, so is promoting post-primary education for girls through fiscal incentives. Interventions that genderize post basic education such as modernizing agriculture at the graduate level is also considered successful.
teaching materials and supplies	Not enough studies (evidence gap)				
community norms and parental attitudes supporting girls' education	Not enough studies (evidence gap): Most studies were multicomponent, making it difficult to establish whether this component was driving effects. The small number of studies that more directly assessed this component had mixed results.		Increasing parental and community engagement can improve enrollment and learning, but programs often come as part of a larger package of interventions so the impact of engagement is harder to disentangle from the other components.	Involving women in school governance and community mobilization has a promising impact on girls' school participation. Gender mainstreaming efforts to change institutional culture is also promising. Women's literacy programs show promising effect on gender norms and identity.	Reducing cultural and social constraints to girls' education is proven successful.

Reviews	Psaki et al. (forthcoming)	Evans and Yuan (2021)	Sperling et al. (2016)	Unterhalter et al. (2014)	Tembon and Fort (2008)
gender sensitivity in the school environment (teachers interact more often with boys and have lower expectations of girls; presence of female teachers)	Not enough studies (evidence gap)		Girl-friendly schools, often multi-component in nature, could improve enrollment and attendance, but more research is need to to determine which components are effective.	Some promising evidence shows that employing female teachers improve girls’ outcomes, but more research is needed to study a wider range of contexts.	Developing and disseminating “gender-sensitive school and pedagogy models” is included in the list of interventions proven successful.
ECD and preschools			Rigorous evidence is “scarce” but some studies suggest positive effects.		
teachers and teaching, including academic support for disadvantaged students and those lagging behind	Effective	Effective in improving girls’ learning (teachers teaching at the right level; structured pedagogy: literacy intervention, mother tongue instruction, and tutor software)	More and better teachers are effective in improving learning outcomes.	Teacher training in “subject content, pedagogy and management” has a strong impact on reducing girls’ dropout. Some promising impact of formal and informal teacher training in gender equality and pedagogy on girls’ learning outcomes. Group learning and learning outside the classroom (tutoring, after-school clubs) has a strong positive effect on learning outcomes	

References

- Abreh, M. K., Agbevanu, W. K., Alhassan, A. J., Ansah, F., Bosu, R. S., Crawford, L., ... Nyame, G. (2021). What happened to dropout rates after COVID-19 school closures in Ghana? Retrieved from Center for Global Development Blog Post website: <https://www.cgdev.org/blog/what-happened-dropout-rates-after-covid-19-school-closures-ghana>
- Adukia, A. (2017). Sanitation and Education. *American Economic Journal: Applied Economics*, 9(2), 23–59. <https://doi.org/10.1257/app.20150083>
- Akresh, R., Halim, D., and Kleemans, M. (2019). *Long-term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia*. Retrieved from https://drive.google.com/file/d/1c7bF9so7c4bokYH5-_epR5muNjpj4Pk7/view
- Alderman, H., and Bundy, D. (2012). School Feeding Programs and Development: Are We Framing the Question Correctly? *The World Bank Research Observer*, 27(2), 204–221. <https://doi.org/10.1093/wbro/lkr005>
- Alexander, J., Kwauk, C., and Perlman-Robinson, J. (2016). *Room to Read: Scaling up literacy through localized solutions across Asia and Africa*. Center for Universal Education at Brookings. Retrieved from Center for Universal Education at Brookings website: <https://www.brookings.edu/wp-content/uploads/2016/07/FINAL-Room-to-Read-Case-Study.pdf>
- Anderson, J. (2018). We know how to make programs that help poor kids. The hard part is scaling up. Retrieved January 30, 2021, from Quartz website: <https://qz.com/1269399/we-know-how-to-make-programs-that-help-poor-kids-the-hard-part-is-scaling-up/>
- Andrabi, T., Das, J., and Khwaja, A. I. (2017). Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets. *American Economic Review*, 107(6), 1535–1563. <https://doi.org/10.1257/aer.20140774>
- Angrist, J., Bettinger, E., and Kremer, M. (2006). Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. *American Economic Review*, 96(3), 847–862. <https://doi.org/10.1257/aer.96.3.847>
- Angrist, N., Evans, D., Filmer, D., Glennerster, R., Rogers, H., and Sabarwal, S. (2020). *How to Improve Education Outcomes Most Efficiently? A Comparison of 150 Interventions Using the New Learning-Adjusted Years of Schooling Metric*. Center for Global Development Working Paper. Retrieved from <https://www.cgdev.org/publication/how-improve-education-outcomes-most-efficiently-comparison-150-interventions-using-new>
- Ardington, C., Wills, G., and Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, 86, 102480. <https://doi.org/10.1016/j.ijedudev.2021.102480>
- ASER. (2021). ASER Centre: Evidence for action. Retrieved from <https://www.asercentre.org/>
- Aurino, E., Gelli, A., Adamba, C., Osei-Akoto, I., and Alderman, H. (2020). Food for thought? Experimental evidence on the learning impacts of a large-scale school feeding program. *Journal of Human Resources*, 1019. <https://doi.org/10.3368/jhr.58.3.1019-10515R1>
- Awason, A., Mpofu, M. E., Mutanguha, M., Muyatwa, M., and Sihwa, P. L. (2020). *Co-Impact Research Proposal- Gender Equitable Educational Outcomes*. Global Office Consulting.

- Bagby, E., Dumitrescu, A., Orfield, C., and Sloan, M. (2016). Long-Term Evaluation of the IMAGINE Project in Niger (English version). *Mathematica Policy Research*. Retrieved from <https://www.mathematica.org/our-publications-and-findings/publications/imagine-report-english-june-2016>
- Baird, S., Ferreira, F. H. G., Özler, B., and Woolcock, M. (2014). Conditional, unconditional and everything in between: A systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6(1), 1–43. <https://doi.org/10.1080/19439342.2014.890362>
- Bajpai, N., Biberman, J., and Sharma, A. (2019). *Information and Communications Technology in the Education Sector in India*. <https://doi.org/10.7916/d8-9anm-mc81>
- Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., and Sulaiman, M. (2020a). Women’s Empowerment in Action: Evidence from a Randomized Control Trial in Africa. *American Economic Journal: Applied Economics*, 12(1), 210–259. <https://doi.org/10.1257/app.20170416>
- Bandiera, O., Buehren, N., Goldstein, M., Rasul, I., and Smurra, A. (2020b). *Do School Closures During an Epidemic have Persistent Effects? Evidence from Sierra Leone in the Time of Ebola*. World Bank Working Paper. Retrieved from http://www.homepages.ucl.ac.uk/~uctpimr/research/ELA_SL.pdf
- Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., ... 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>
- Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., and Khemani, S. (2010). Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India. *American Economic Journal: Economic Policy*, 2(1), 1–30. <https://doi.org/10.1257/pol.2.1.1>
- Banerjee, A. V., Cole, S., Duflo, E., and Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics*, 122(3), 1235–1264. <https://doi.org/10.1162/qjec.122.3.1235>
- Banerji, R., Berry, J., and Shotland, M. (2017). The Impact of Maternal Literacy and Participation Programs: Evidence from a Randomized Evaluation in India. *American Economic Journal: Applied Economics*, 9(4), 303–337. <https://doi.org/10.1257/app.20150390>
- Baric, S. (2013). Where the boys are: Engaging young adolescent boys in support of girls’ education and leadership. *Gender & Development*, 21(1), 147–160. <https://doi.org/10.1080/13552074.2013.769777>
- Barr, A., Bategeka, L., Guloba, M., Kasirye, I., Mugisha, F., Serneels, P., and Zeitlin, A. (2012). Management and Motivation in Ugandan Primary Schools: An impact evaluation report. In *Working Papers PIERI* (No. 2012–14). PEP-PIERI. Retrieved from PEP-PIERI website: <https://ideas.repec.org/p/lvl/piercr/2012-14.html>
- Bates, M. A., and Glennerster, R. (2017). The Generalizability Puzzle. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/the_generalizability_puzzle
- Benhassine, N., Devoto, F., Duflo, E., Dupas, P., and Pouliquen, V. (2015). Turning a Shove into a Nudge? A “Labeled Cash Transfer” for Education. *American Economic Journal: Economic Policy*, 7(3), 86–125. <https://doi.org/10.1257/pol.20130225>

- Benshaul-Tolonen, A., Aguilar-Gomez, S., Batzer, N. H., Cai, R., and Nyanza, E. C. (2020a). Period teasing, stigma and knowledge: A survey of adolescent boys and girls in Northern Tanzania. *PLOS ONE*, *15*(10), e0239914. <https://doi.org/10.1371/journal.pone.0239914>
- Benshaul-Tolonen, A., Zulaika, G., Nyothach, E., Odour, C., Mason, L., Obor, D., ... Phillips-Howard, P. (2021). *Sanitary products, absenteeism and psychosocial well-being: Evidence from a three-arm cluster randomized controlled feasibility study in Western Kenya* [Working Paper]. CDEP-CGEG Working Paper Series No. 93. Retrieved from CDEP-CGEG Working Paper Series No. 93 website: <https://anjatolonen.files.wordpress.com/2021/03/wp93tolonen.pdf>
- Benshaul-Tolonen, A., Zulaika, G., Sommer, M., and Phillips-Howard, P. A. (2020b). Measuring menstruation-related absenteeism among adolescents in low-income countries. In C. Bobel, I. T. Winkler, B. Fahs, K. A. Hasson, E. A. Kissling, and T.-A. Roberts (Eds.), *The Palgrave Handbook of Critical Menstruation Studies* (pp. 705–723). Singapore: Springer. https://doi.org/10.1007/978-981-15-0614-7_52
- Berlinski, S., Busso, M., Dinkelman, T., and Martinez, C. (2016). *Reducing Parent-School Information Gaps and Improving Education Outcomes: Evidence from High Frequency Text Messaging in Chile*. Retrieved from https://assets.publishing.service.gov.uk/media/58778bfa40f0b60e4c000135/726__Reducing-Parent-School-information-gap_BBDM-Dec2016.pdf
- Blimpo, M. P., Gajigo, O., and Pugatch, T. (2019). Financial Constraints and Girls' Secondary Education: Evidence from School Fee Elimination in The Gambia. *The World Bank Economic Review*, *33*(1), 185–208. <https://doi.org/10.1093/wber/lhw051>
- Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., and Sandefur, J. (2018). Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics*, *168*, 1–20. <https://doi.org/10.1016/j.jpubeco.2018.08.007>
- Buehren, N., Goldstein, M. P., Gulesci, S., Sulaiman, M., and Yam, V. (2017). *Evaluation of an adolescent development program for girls in Tanzania* (No. WPS7961; pp. 1–25). The World Bank. Retrieved from The World Bank website: <http://documents.worldbank.org/curated/en/245071486474542369/Evaluation-of-an-adolescent-development-program-for-girls-in-Tanzania>
- Bulman, G., and Fairlie, R. W. (2016). Chapter 5: Technology and Education: Computers, Software, and the Internet. In *Handbook of the Economics of Education, Volume 5*. Elsevier B.V. Retrieved from https://people.ucsc.edu/~gbulman/tech_2016.pdf
- Cameron, D. B., Mishra, A., and Brown, A. N. (2016). The growth of impact evaluation for international development: How much have we learned? *Journal of Development Effectiveness*, *8*(1), 1–21. <https://doi.org/10.1080/19439342.2015.1034156>
- Carvalho, S., Evans, D., and Mendez Acosta, A. (2021). *Beyond Average: How to Reach Girls who Face Multiple Sources of Exclusion*. Center for Global Development.
- Chakraborty, T., and Jayaraman, R. (2019). School feeding and learning achievement: Evidence from India's midday meal program. *Journal of Development Economics*, *139*, 249–265. <https://doi.org/10.1016/j.jdeveco.2018.10.011>
- Chelwa, G., Pellicer, M., and Maboche, M. (2019). Teacher Pay and Educational Outcomes: Evidence from the Rural Hardship Allowance in Zambia. *South African Journal of Economics*, *87*(3), 255–282. <https://doi.org/10.1111/saje.12227>

- Chemonics International Inc. (2019). *Sindh Reading Program: Final Report*. Retrieved from https://chemonics.com/wp-content/uploads/2019/08/Pakistan-SRP_FinalReport_2019.pdf
- Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, N., Taylor, S., and Thulare, T. (2021). *Can Virtual Replace In-person Coaching? Experimental Evidence on Teacher Professional Development and Student Learning in South Africa* [RISE Working Paper 20/050.]. Retrieved from <https://riseprogramme.org/publications/can-virtual-replace-person-coaching-experimental-evidence-teacher-professional>
- Cilliers, J., Fleisch, B., Prinsloo, C., and Taylor, S. (2019). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources*, 56(3), 0618-9538R1. <https://doi.org/10.3368/jhr.55.3.0618-9538R1>
- Cilliers, J., Mbiti, I. M., and Zeitlin, A. (2020). Can Public Rankings Improve School Performance? Evidence from a Nationwide Reform in Tanzania. *Journal of Human Resources*, 0119-9969R1. <https://doi.org/10.3368/jhr.56.3.0119-9969R1>
- Crawford, L. (2020). Why the COVID Crisis Is Not Edtech's Moment in Africa. Retrieved March 10, 2021, from Center For Global Development Blog Post website: <https://www.cgdev.org/blog/why-covid-crisis-not-edtechs-moment-africa>
- Crawford, L., and Hares, S. (2020). Introducing the Girl's Education Policy Index. Retrieved from <https://www.cgdev.org/blog/introducing-girls-education-policy-index>
- Crawford, L., Hares, S., and Minardi, A. L. (2021). New data on learning loss in Pakistan. Retrieved from Center For Global Development Blog Post website: <https://www.cgdev.org/blog/new-data-learning-loss-pakistan>
- Cristia, J., Czerwonko, A., and Garofalo, P. (2014). Does Technology in Schools Affect Repetition, Dropout and Enrollment? Evidence from Peru. *Journal of Applied Economics*, 17(1), 89–111. [https://doi.org/10.1016/S1514-0326\(14\)60004-0](https://doi.org/10.1016/S1514-0326(14)60004-0)
- Cristia, J., Ibararán, P., Cueto, S., Santiago, A., and 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>
- Cull, B., and McKenzie, D. (2020, March). Implementing successful small interventions at a large scale is hard. Retrieved from World Bank Development Impact Blogs website: <https://blogs.worldbank.org/developmenttalk/implementing-successful-small-interventions-large-scale-hard>
- Davis, M., Ingwersen, N., Kazianga, H., Linden, L., Mamun, A., Protik, A., and Sloan, M. (2016). *Ten-Year Impacts of Burkina Faso's BRIGHT Program*. Mathematica Policy Research. Retrieved from Mathematica Policy Research website: <https://www.mathematica.org/our-publications-and-findings/publications/ten-year-impacts-of-burkina-fasos-bright-program>
- DellaVigna, S., and Linos, E. (2021). *RCTs to Scale: Comprehensive Evidence from Two Nudge Units*. Working Paper, Cambridge, MA. <https://doi.org/10.3386/w27594>
- Deschênes, S., and Hotte, R. (2019). *Assessing the Effects of an Education Policy on Women's Well-being: Evidence from Benin* (No. halshs-02179704). HAL. Retrieved from HAL website: <https://ideas.repec.org/p/hal/psewpa/halshs-02179704.html>
- Dhar, D., Jain, T., and Jayachandran, S. (2020). *Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India*. Retrieved from https://faculty.wcas.northwestern.edu/~sjv340/reshaping_gender_attitudes.pdf

- Dhar, V., Nilekani, N., Maruwada, S., and Pappu, N. (2016). Big Data As an Enabler of Primary Education. *Big Data*, 4(3), 137–140. <https://doi.org/10.1089/big.2016.29011.vdc>
- Drake, L., Fernandes, M., Aurino, E., Kiamba, J., Giyose, B., Burbano, C., ... Gelli, A. (2017). School Feeding Programs in Middle Childhood and Adolescence. In D. A. P. Bundy, N. de Silva, S. Horton, D. T. Jamison, and G. C. Patton (Eds.), *Child and Adolescent Health and Development* (3rd ed.). Washington (DC): The International Bank for Reconstruction and Development / The World Bank. Retrieved from <http://www.ncbi.nlm.nih.gov/books/NBK525249/>
- Duflo, A., Kiessel, J., and Lucas, A. (2020). *External Validity: Four Models of Improving Student Achievement* (Working Paper No. 27298). National Bureau of Economic Research. <https://doi.org/10.3386/w27298>
- Duflo, E., Dupas, P., and Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools. *Journal of Public Economics*, 123, 92–110. <https://doi.org/10.1016/j.jpubeco.2014.11.008>
- Duflo, E., Dupas, P., and Kremer, M. (2021). The Impact of Free Secondary Education: Experimental Evidence from Ghana. *NBER Working Paper*. Retrieved from <https://www.nber.org/papers/w28937>
- Duflo, E., Hanna, R., and 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>
- Edmonds, E., Feigenberg, B., and Leight, J. (2019). *Advancing the Agency of Adolescent Girls: Final Evaluation Report for Room to Read*. Retrieved from <https://www.roomtoread.org/media/gkni2y0u/gep-evaluation-2016-18-rajasthan-final-report.pdf>
- Escueta, M., Nickow, A. J., Oreopoulos, P., and Quan, V. (2020). Upgrading Education with Technology: Insights from Experimental Research. *Journal of Economic Literature*, 58(4), 897–996. <https://doi.org/10.1257/jel.20191507>
- Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research*, 49(4), 343–363. <https://doi.org/10.1080/00131880701717198>
- Evans, A. (2019). How Cities Erode Gender Inequality: A New Theory and Evidence from Cambodia. *Gender & Society*, 33(6), 961–984. <https://doi.org/10.1177/0891243219865510>
- Evans, D. K., Akmal, M., and Jakiela, P. (2021a). Gender Gaps in Education: The Long View. *IZA Journal of Development and Migration*.
- Evans, D. K., Gale, C., and Kosec, K. (2020). *The Education Impacts of Cash Transfers for Children with Multiple Indicators of Vulnerability*. Working Paper, Center for Global Development. Retrieved from <https://www.cgdev.org/publication/education-impacts-cash-transfers-children-multiple-indicators-vulnerability>
- Evans, D. K., Hares, S., Mendez Acosta, A., and Saintis, C. (2021b). It's Been a Year Since Schools Started to Close Due to COVID-19. Retrieved March 8, 2021, from Center For Global Development Blog Post website: <https://www.cgdev.org/blog/its-been-year-schools-started-close-due-covid-19>
- Evans, D. K., and Mendez Acosta, A. (2021). Education in Africa: What Are We Learning? *Journal of African Economies*, 30(1), 13–54. <https://doi.org/10.1093/jae/ejaa009>

- Evans, D. K., and Popova, A. (2016). Cost-Effectiveness Analysis in Development: Accounting for Local Costs and Noisy Impacts. *World Development*, 77, 262–276. <https://doi.org/10.1016/j.worlddev.2015.08.020>
- Evans, D. K., and Yuan, F. (2019). *Equivalent years of schooling: A metric to communicate learning gains in concrete terms* (No. WPS8752; pp. 1–52). The World Bank. Retrieved from The World Bank website: <http://documents.worldbank.org/curated/en/123371550594320297/Equivalent-Years-of-Schooling-A-Metric-to-Communicate-Learning-Gains-in-Concrete-Terms>
- Evans, D. K., and Yuan, F. (2020). *How Big Are Effect Sizes in International Education Studies?* Center for Global Development. Retrieved from <https://www.cgdev.org/publication/how-big-are-effect-sizes-international-education-studies>
- Evans, D. K., and Yuan, F. (2021). What we learn about girls’ education from interventions that do not focus on girls. *The World Bank Economic Review*, (lhab007). <https://doi.org/10.1093/wber/lhab007>
- Faul, M. (2016). *Scoping Study: Education Systems Diagnostics Tools*. Retrieved from https://www.academia.edu/31407684/Scoping_Study_Education_Systems_Diagnostics_Tools
- Fazzio, I., Eble, A., Lumsdaine, R., Boone, P., Bouy, B., Hsieh, J., ... Silva, A. F. (2020). *Large Learning Gains in Pockets of Extreme Poverty: Experimental Evidence from Guinea Bissau*. Working Paper. Retrieved from <http://www.alexeble.com/research>
- Fiala, N., Garcia-Hernandez, A., Narula, K., and Prakash, N. (2020, January). *Wheels of Change: Transforming Women’s Lives with Education and Empowerment*. Presented at the Center for the Study of African Economies Conference Papers. Retrieved from https://www.dropbox.com/s/is9sc8khliipbx3/Slides_Cycles%20%28Dec%202019%29.pdf?dl=0
- Fraudenberger, E., and Davis, J. (2017). *Tusome External Evaluation Midline Report*. Retrieved from <https://ierc-publicfiles.s3.amazonaws.com/public/resources/Tusome%20Midline%20evaluation%202017%20final%20report%20from%20DEC.pdf>
- Gelli, A., Meir, U., and Espejo, F. (2007). Does Provision of Food in School Increase Girls’ Enrollment? Evidence from Schools in Sub-Saharan Africa. *Food and Nutrition Bulletin*, 28(2), 149–155. <https://doi.org/10.1177/156482650702800203>
- Glewwe, P., Park, A., and Zhao, M. (2016). A better vision for development: Eyeglasses and academic performance in rural primary schools in China. *Journal of Development Economics*, 122, 170–182. <https://doi.org/10.1016/j.jdeveco.2016.05.007>
- Global Education Evidence Advisory Panel. (2020). *Cost-Effective Approaches to Improve Global Learning: What Does Recent Evidence Tell Us Are “Smart Buys” for Improving Learning in Low and Middle Income Countries?* World Bank. Retrieved from <http://documents1.worldbank.org/curated/en/719211603835247448/pdf/Cost-Effective-Approaches-to-Improve-Global-Learning-What-Does-Recent-Evidence-Tell-Us-Are-Smart-Buys-for-Improving-Learning-in-Low-and-Middle-Income-Countries.pdf>
- Goertz, G., and Mahoney, J. (2012). Causes-of-Effects versus Effects-of-Causes. In *A Tale of Two Cultures*. Princeton University Press. <https://doi.org/10.23943/princeton/9780691149707.003.0003>
- Gove, A., Poole, M. K., and Piper, B. (2017). Designing for Scale: Reflections on Rolling Out Reading Improvement in Kenya and Liberia. *New Directions for Child and Adolescent Development*, 2017(155), 77–95. <https://doi.org/10.1002/cad.20195>

- Hartmann, A., and Linn, J. (2008). *Scaling Up: A framework and lessons for development effectiveness from literature and practice* [Wolfensohn Center for Development Working Paper 5]. Brookings. Retrieved from Brookings website: https://www.brookings.edu/wp-content/uploads/2016/06/10_scaling_up_aid_linn.pdf
- Islam, A. (2019). Parent–teacher meetings and student outcomes: Evidence from a developing country. *European Economic Review*, 111, 273–304. <https://doi.org/10.1016/j.euroecorev.2018.09.008>
- Jackson, C. K. (2019). Can Introducing Single-Sex Education into Low-Performing Schools Improve Academics, Arrests, and Teen Motherhood? *Journal of Human Resources*, 0618. <https://doi.org/10.3368/jhr.56.1.0618-9558R2>
- Joddar, P. (2018). *Impact Evaluation of the Literacy Program- Partnership Approach under Scaling up Early Reading Intervention(SERI) funded by USAID: 2018 Endline Evaluation Report for Chhattisgarh and Uttarakhand*.
- Kazianga, H., Levy, D., Linden, L. L., and Sloan, M. (2013). The Effects of “Girl-Friendly” Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso. *American Economic Journal: Applied Economics*, 5(3), 41–62. <https://doi.org/10.1257/app.5.3.41>
- Kerwin, J. T., and Thornton, R. L. (2021). Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures. *The Review of Economics and Statistics*, 103(2), 251–264. https://doi.org/10.1162/rest_a_00911
- Kim, Jim Y. (2016). *World Bank Group President Jim Yong Kim's Remarks at “Let Girls’ Learn” Spring Meetings Event*. Text/HTML presented at the Washington DC, United States. Washington DC, United States. Retrieved from <https://www.worldbank.org/en/news/speech/2016/04/13/world-bank-group-president-jim-yong-kims-remarks-at-let-girls-learn-spring-meetings-event>
- Kim, J., Rose, P., Tiruneh, D. T., Sabates, R., and Woldehanna, T. (2021). Learning Inequalities Widen Following COVID-19 School Closures in Ethiopia. Retrieved from Center For Global Development Blog Post website: <https://riseprogramme.org/blog/learning-inequalities-widen-COVID-19-Ethiopia>
- Kraft, M. A., and Blazar, D. (2018, July 31). Taking Teacher Coaching To Scale. Retrieved January 30, 2021, from Education Next website: <https://www.educationnext.org/taking-teacher-coaching-to-scale-can-personalized-training-become-standard-practice/>
- Kwauk, C., and Robinson, J. P. (2016). *Sistema de Aprendizaje Tutorial: Redefining rural secondary education in Latin America*. The Brookings Institution. Retrieved from The Brookings Institution website: <https://www.brookings.edu/wp-content/uploads/2016/07/FINAL-SAT-Case-Study.pdf>
- Kwauk, C., Schmidt, D., and Ganju, E. (2021). What do we know about the effects of COVID-19 on girls’ return to school? Retrieved from Brookings Education Plus Development Blog website: <https://www.brookings.edu/blog/education-plus-development/2021/09/22/what-do-we-know-about-the-effects-of-covid-19-on-girls-return-to-school/>
- Lichand, G., Alberto Doria, C., Leal Neto, O., and Cossi, J. (2021). *The Impacts of Remote Learning in Secondary Education: Evidence from Brazil during the Pandemic* (SSRN Scholarly Paper No. ID 3841775). Rochester, NY: Social Science Research Network. <https://doi.org/10.2139/ssrn.3841775>

- Linden, L., and Barrera-Osorio, F. (2009). *The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia*. Retrieved from <https://openknowledge.worldbank.org/bitstream/handle/10986/4032/WPS4836.pdf?sequence=1>
- London, J. D. (2021). *Outlier Vietnam and the Problem of Embeddedness: Contributions to the Political Economy of Learning*. RISE Working Paper 21/062. Retrieved from RISE Working Paper 21/062 website: <https://riseprogramme.org/publications/outlier-vietnam-and-problem-embeddedness-contributions-political-economy-learning>
- Loureiro, A., Cruz, L., Lautharte, I., and Evans, D. (2020). *The State of Ceara in Brazil is a Role Model for Reducing Learning Poverty*. World Bank. Retrieved from World Bank website: <https://openknowledge.worldbank.org/handle/10986/34156>
- Loyalka, P., Popova, A., Li, G., and Shi, Z. (2019). Does Teacher Training Actually Work? Evidence from a Large-Scale Randomized Evaluation of a National Teacher Training Program. *American Economic Journal: Applied Economics*, 11(3), 128–154. <https://doi.org/10.1257/app.20170226>
- Management Systems International. (2018). *2017 Early Grade Reading Assessment: Sindh-Pakistan Reading Program*. Retrieved from <https://earlygradereadingbarometer.org/files/Pakistan%20Sindh%20PRP.pdf>
- Mares, M.-L., and Pan, Z. (2013). Effects of Sesame Street: A meta-analysis of children's learning in 15 countries. *Journal of Applied Developmental Psychology*, 34(3), 140–151. <https://doi.org/10.1016/j.appdev.2013.01.001>
- Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., and Rajani, R. (2019). 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>
- McEwan, P. J. (2013). The impact of Chile's school feeding program on education outcomes. *Economics of Education Review*, 32, 122–139. <https://doi.org/10.1016/j.econedurev.2012.08.006>
- 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>
- McEwan, P. J., Murphy-Graham, E., Torres Iribarra, D., Aguilar, C., and Rápalo, R. (2015). Improving Middle School Quality in Poor Countries: Evidence From the Honduran Sistema de Aprendizaje Tutorial. *Educational Evaluation and Policy Analysis*, 37(1), 113–137. <https://doi.org/10.3102/0162373714527786>
- Megevand, M., and Marchesini, L. (2020). *Women and Girls Safe Spaces: A Toolkit for Advancing Women's and Girls' Empowerment in Humanitarian Settings*. International Rescue Committee and International Medical Corps. Retrieved from International Rescue Committee and International Medical Corps website: <https://gbvaor.net/sites/default/files/2020-02/IRC-WGSS-Toolkit-Eng.pdf>
- Mendez Acosta, A., and Evans, D. (2020). COVID-19 and Girls' Education: What We Know So Far and What We Expect. Retrieved March 8, 2021, from Center For Global Development website: <https://www.cgdev.org/blog/covid-19-and-girls-education-what-we-know-so-far-and-what-we-expect-happen>

- Miske Witt & Associates Inc. (2011). *The Power to Lead Alliance (PTLA) Empowering Girls to Learn and Lead: Final Evaluation Report for CARE USA*. Retrieved from http://careevaluations.org/wp-content/uploads/PTLA-FINAL-EVALUATION-REPORT_MWAI.pdf
- Muralidharan, K., and Prakash, N. (2017). Cycling to School: Increasing Secondary School Enrollment for Girls in India. *American Economic Journal: Applied Economics*, 9(3), 321–350. <https://doi.org/10.1257/app.20160004>
- Muralidharan, K., and Singh, A. (2021, September). *Improving Schooling Productivity Through Computer-Aided Instruction: Experimental Evidence from Rajasthan*. RISE Conference Paper 2021.
- Muralidharan, K., Singh, A., and Ganimian, A. J. (2019). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. *American Economic Review*, 109(4), 1426–1460. <https://doi.org/10.1257/aer.20171112>
- Muralidharan, K., and Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119(1), 39–77. <https://doi.org/10.1086/659655>
- Pandey, P., Goyal, S., and Sundararaman, V. (2008). *Community Participation in Public Schools: The Impact of Information Campaigns in Three Indian States*. Retrieved from <https://openknowledge.worldbank.org/handle/10986/6359>
- Pappas, G., Agha, A., Rafique, G., Khan, K. S., Badruddin, S., and Peermohamed, H. (2008). Community-based approaches to combating malnutrition and poor education among girls in resource-poor settings: Report of a large scale intervention in Pakistan. *Rural and Remote Health*, 8(1), 820.
- Perera, M., and Aboal, D. (2019). *The impact of a mathematics computer-assisted learning platform on students' mathematics test scores*. Retrieved from </paper/The-impact-of-a-mathematics-computer-assisted-on-Perera-Aboal/ae058393d4f9da5fa415acde66fed4a31271a08e>
- Piper, B., Destefano, J., Kinyanjui, E. M., and Ong'ele, S. (2018). Scaling up successfully: Lessons from Kenya's Tusome national literacy program. *Journal of Educational Change*, 19(3), 293–321. <https://doi.org/10.1007/s10833-018-9325-4>
- Piper, B., and Mugenda, A. (2014). *The Primary Math and Reading (PRIMR) Initiative: Endline Impact Evaluation*. USAID. Retrieved from USAID website: https://pdf.usaid.gov/pdf_docs/PA00K27S.pdf
- Popova, A., Evans, D. K., Breeding, M. E., and Arancibia, V. (2021). Teacher Professional Development around the World: The Gap between Evidence and Practice. *World Bank Research Observer*.
- Presidential Policy and Strategy Unit Kenya, and Population Council. (2021). *Promises to keep: Impact of COVID-19 on adolescents in Kenya*. Nairobi: Population Council. Retrieved from Population Council website: https://knowledgecommons.popcouncil.org/cgi/viewcontent.cgi?article=2378&context=departments_sbsr-pgy&mc_cid=8d27443b6f&mc_eid=a25e001d8c
- Pritchett, L., and Sandefur, J. (2013). Context Matters for Size: Why External Validity Claims and Development Practice do not Mix. *Journal of Globalization and Development*, 4(2), 161–197. <https://doi.org/10.1515/jgd-2014-0004>

- Pritchett, L., and Sandefur, J. (2015). Learning from Experiments When Context Matters. *American Economic Review*, 105(5), 471–475. <https://doi.org/10.1257/aer.p20151016>
- Psaki, S. R., Haberland, N. A., Kozak, M. L., and Woyczynski, L. (2021). *Girls' Education Roadmap: 2021 Report* [EGER Reports]. New York: Population Council. Retrieved from Population Council website: https://egeresource.org/documents/4/Girls_Education_Roadmap_2021_Report.pdf
- Psaki, S. R., Haberland, N. A., Mensch, B. S., Woyczynski, L., and Chuang, E. K. (forthcoming). Policies and interventions to remove gender-related barriers to girls' school participation and learning in low- and middle-income countries: A systematic review of the evidence. *Campbell Systematic Reviews*. <https://doi.org/10.1002/cl2.1047>
- Pugatch, T., and Schroeder, E. (2018). Teacher pay and student performance: Evidence from the Gambian hardship allowance. *Journal of Development Effectiveness*, 10(2), 249–276. <https://doi.org/10.1080/19439342.2018.1452778>
- Randall, J., and Garcia, A. (2020). Let's Go Girls!: Evaluating the Effectiveness of Tutoring and Scholarships on Primary School Girls' Attendance and Academic Performance in the Democratic Republic of the Congo. *FIRE: Forum for International Research in Education*, 6(3), 19–35.
- Read, L., and Atinc, T. M. (2018). Information for accountability: Transparency and citizen engagement for improved service delivery in education systems. *Journal of International Cooperation in Education*, 20(2/21–2), 35–65.
- Reinikka, R., and Svensson, J. (2011). The power of information in public services: Evidence from education in Uganda. *Journal of Public Economics*, 95(7), 956–966. <https://doi.org/10.1016/j.jpubeco.2011.02.006>
- Ripley, A. (2014). *The Smartest Kids in the World: And How They Got That Way* (Reprint edition). New York: Simon & Schuster.
- Robinson, J. P., Hannahan, P., and Curtiss, M. (2019, February 8). Millions Learning. Retrieved January 30, 2021, from Brookings website: <https://www.brookings.edu/product/millions-learning/>
- Rodriguez-Segura, D. (2021). EdTech in Developing Countries: A Review of the Evidence. *The World Bank Research Observer*, (lkab011). <https://doi.org/10.1093/wbro/lkab011>
- Ruddle, N., and Rawle, G. (2020). *EQUIP-Tanzania Impact Evaluation: Final Endline Report*. Retrieved from https://www.opml.co.uk/files/Publications/8383-assessing-equip-t/1589895669_opm-ie-final-endline-report-final.pdf?noredirect=1
- Sabates, R., Rose, P., Alcott, B., and Delprato, M. (2020). Assessing cost-effectiveness with equity of a programme targeting marginalised girls in secondary schools in Tanzania. *Journal of Development Effectiveness*, 0(0), 1–19. <https://doi.org/10.1080/19439342.2020.1844782>
- Sabet, S. M., and Brown, A. N. (2018). Is impact evaluation still on the rise? The new trends in 2010–2015. *Journal of Development Effectiveness*, 10(3), 291–304. <https://doi.org/10.1080/19439342.2018.1483414>
- Sahlberg, P., Robinson, K., and Gardner, H. (2021). *Finnish Lessons 3.0: What Can the World Learn from Educational Change in Finland?* (3rd edition). New York: Teachers College Press.
- Savage, L., and Martinez, R. (2019). How Can We Best Diagnose Education Systems to Improve Them. Retrieved from Global Partnership for Education website: <https://www.globalpartnership.org/blog/how-can-we-best-diagnose-education-systems-improve-them>

- Schurmann, A. T. (2009). Review of the Bangladesh Female Secondary School Stipend Project Using a Social Exclusion Framework. *Journal of Health, Population and Nutrition*, 27(4), 505–517.
- Snilstveit, B., Stevenson, J., Menon, R., Philips, D., Gallagher, E., Geelen, M., ... Jimenez, E. (2016). *The impact of education programmes on learning and school participation in low- and middle-income countries*. London: International Initiative for Impact Evaluation (3ie). Retrieved from International Initiative for Impact Evaluation (3ie) website: <https://www.3ieimpact.org/sites/default/files/2019-05/srs7-education-report.pdf>
- Sperling, G., Winthrop, R., and Kwauk, C. (2016). *What Works in Girls' Education: Evidence for the World's Best Investment*. Brookings Institution Press. Retrieved from <https://www.jstor.org/stable/10.7864/j.ctt17w8hh8>
- Stern, J., Jukes, M., and Piper, B. (2020). Is It Possible to Improve Learning at Scale? Reflections on the Process of Identifying Large-Scale Successful Education Interventions. Retrieved March 10, 2021, from Center For Global Development website: <https://www.cgdev.org/blog/it-possible-improve-learning-scale-reflections-process-identifying-large-scale-successful>
- Subrahmanian, R. (2005). *Scaling up good practices in girls' education*. UNESCO. Retrieved from UNESCO website: <https://unesdoc.unesco.org/ark:/48223/pf0000142240?posInSet=3&queryId=c1b3bcb2-ef1d-4c3b-baef-49f954dbbbbf>
- Tembon, M., and Fort, L. (2008). *Girls' Education in the 21st Century: Gender Equality, Empowerment, and Economic Growth*. World Bank. Retrieved from World Bank website: <https://openknowledge.worldbank.org/bitstream/handle/10986/6554/454040PUB097801E0ONLY10Aug012102008.pdf?sequence=1&isAllowed=y>
- Tricco, A. C., Langlois, E. V., and Straus, S. E. (2017). *Rapid reviews to strengthen health policy and systems: A practical guide*. World Health Organization. Retrieved from World Health Organization website: <http://apps.who.int/iris/bitstream/handle/10665/258698/9789241512763-eng.pdf;jsessionid=14116357F6DDA12F8E02EBDCFC5F2D59?sequence=1>
- United Nations. (2015). Sustainable Development Goal 5: Achieve Gender Equality and Empower All Women and Girls. Retrieved January 16, 2021, from <https://sdgs.un.org/goals/goal5>
- Unterhalter, E., North, A., Arnot, M., Lloyd, C., Moletsane, L., Murphy-Graham, E., ... Saito, M. (2014). *Interventions to enhance girls' education and gender equality*. Department for International Development. Retrieved from Department for International Development website: <https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/Girls%20education%202014%20Unterhalter%20report.pdf?ver=2015-12-08-165815-117>
- Uwezo. (2021). Uwezo Kenya. Retrieved from <https://www.uwezo.net/>
- Williams, T. P., Abbott, P., and Mupenzi, A. (2015). 'Education at our school is not free': The hidden costs of fee-free schooling in Rwanda. *Compare: A Journal of Comparative and International Education*, 45(6), 931–952. <https://doi.org/10.1080/03057925.2014.938611>
- World Bank. (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Washington, DC. Retrieved from <https://www.worldbank.org/en/publication/wdr2018>
- World Bank. (2020a). *Strategic Impact Evaluation Fund: Presenting the evaluations of the COVID-19 Emergency Window*.

- World Bank. (2020b). *The Human Capital Index 2020 Update: Human Capital in the Time of COVID-19*. Retrieved from <https://openknowledge.worldbank.org/handle/10986/34432>
- World Bank. (2020c). World Development Indicators. Retrieved March 4, 2020, from <https://databank.worldbank.org/source/world-development-indicators#>
- World Bank. (2021). Service Delivery Indicators. Retrieved from <https://www.sdindicators.org/#:~:text=Indicators,a%20school%20or%20health%20facility>.
- Yanguas, M. L. (2020). Technology and educational choices: Evidence from a one-laptop-per-child program. *Economics of Education Review*, 76, 101984. <https://doi.org/10.1016/j.econedurev.2020.101984>