What We Learn about Girls’ Education from Interventions that Do Not Focus on Girls

David K. Evans and Fei Yuan

Abstract

Despite dramatic global gains in access to education, 130 million girls of school age remain out of school. Among those who do enter, too many do not gain the essential skills to succeed after they complete their schooling. Previous efforts to synthesize evidence on how to improve educational outcomes for girls have tended to focus on interventions that are principally targeted to girls, such as girls’ latrines or girls’ scholarships. But if general, non-targeted interventions—those that benefit both girls and boys—significantly improve girls’ education, then focusing only on girl-targeted interventions may miss some of the best investments for improving educational opportunities for girls in absolute terms. This review brings together evidence from 270 educational interventions from 177 studies in 54 low- and middle-income countries and identifies their impacts on girls, regardless of whether the interventions specifically target girls. The review finds that to improve access and learning, general interventions deliver gains for girls that are comparable to girl-targeted interventions. At the same time, many more general interventions have been tested, providing a broader menu of options for policy makers. General interventions have similar impacts for girls as for boys. Many of the most effective interventions to improve access for girls are household-based (such as cash transfer programs), and many of the most effective interventions to improve learning for girls involve improving the pedagogy of teachers. Girl-targeted interventions may make the most sense when addressing constraints that are unique to girls.

Keywords: education, girls, effect size, impact evaluation, economic development

JEL: I21, I24, J16, O1
What We Learn about Girls’ Education from Interventions that Do Not Focus on Girls

David K. Evans
Center for Global Development
devans@cgdev.org

Fei Yuan
Harvard Graduate School of Education
fyuan@g.harvard.edu

The authors would like to thank Deon Filmer, Erin Ganju, Markus Goldstein, Pamela Jakiela, Oni Lusk-Stover, Mary Obelnicki, Owen Ozier, Pauline Rose, Dana Schmidt, Craig Silverstein, Lexie Wagner, Kim Wright-Violicl, Louise Yorke, and various seminar audiences for feedback and suggestions. We would also like to thank Echidna Giving, the Umbrella Facility for Gender Equality at the World Bank, and the Bill & Melinda Gates Foundation for financing this work. We thank Tara Siegel, Danielle Sobol, and Shikhty Sunny for excellent research assistance.

# Contents

1. Introduction.................................................................................................................................. 1

2. Literature review: The missing evidence .................................................................................. 4

3. Method........................................................................................................................................... 4
   Literature search........................................................................................................................... 5
   Inclusion criteria .......................................................................................................................... 5
   Data collection................................................................................................................................ 6
   Coding of effect sizes.................................................................................................................. 7
   Qualitative variables.................................................................................................................... 8

4. Results............................................................................................................................................ 8
   Are girl-targeted interventions the most effective for girls? ................................................. 8
   For general interventions, do impacts on girls tend to be larger?........................................... 10
   What are the most effective interventions for girls? ............................................................ 10

5. Discussion ................................................................................................................................... 12
   Inequality..................................................................................................................................... 12
   Costs............................................................................................................................................ 13
   Program attributes ..................................................................................................................... 13
   What has been studied.............................................................................................................. 14

6. Conclusion................................................................................................................................... 14

References........................................................................................................................................ 15

Figures and tables........................................................................................................................... 19

Appendix A. Studies included ...................................................................................................... 34

Appendix B. Journals and websites searched............................................................................. 46
1. Introduction

Investing in girls’ education has been called “the world’s best investment” (Sperling and Winthrop 2015). Roughly 15 percent of the world’s population in 2016 are girls younger than 18 years old. Educating girls frees them to raise their aspirations and increases their potential to achieve them. It allows them to access better health, higher earnings, and more control over their lives, which subsequently leads to positive impacts on their families, communities, and societies (World Bank 2018b). More and better education for girls and boys contributes to economic growth, labor productivity, and social mobility.

The effects of girls’ education extend across generations. Better educated women tend to have lower fertility rates (Castro Martin 1995). They also have lower child mortality rates: data from 175 countries between 1970 and 2009 show that 51 percent of the reduction in deaths in children younger than 5 years in the past 40 years could be attributed to the increased education attainment in women of reproductive age (Gakidou et al. 2010). In addition, mothers’ education is strongly associated with children’s educational attainment and achievement (Ermisch and Francesconi 2001; Schultz 2002).

Despite all the demonstrated benefits, girls face challenges in education that boys do not. For example, girls’ schooling is more sensitive to both direct costs and opportunity costs than that of boys (King and Winthrop 2015). When girls become teenagers, early marriage and adolescent pregnancy may prevent them pursuing further schooling (Field and Ambrus 2008). Social norms in certain cultural contexts may restrict them from accessing economic opportunities as well, potentially limiting the gains from investing in education. For example, 90 percent of women in Uttar Pradesh, India, reported that they needed their husband’s permission to work (World Bank 2018b). Even when women participate in the labor market, they are less likely to find a job. In the Middle East and North Africa, the female unemployment rate is twice that of their male counterparts (ILO 2017). Educating girls increases their control over their own lives – regardless of whether they participate in the formal labor market – and empowers them to advance gender equality in other areas.

Globally, the gender gap in primary and secondary schooling has significantly narrowed – and in some countries, reversed – over the last two decades (Psaki, McCarthy, and Mensch 2018), but 130 million girls remain out of school (UNESCO 2016 ). However, global averages mask the large variation in gender inequality at the regional and national levels. In many countries, girls are still more likely to be excluded from education (UNESCO 2016 ). Girls also complete less schooling than boys. Gender inequality (favoring boys) in primary completion remains in 60 of all 177 countries with available data (Figure 1A) and is more prominent in poor countries: 21 of the 25 countries with a gap larger than 5 percent are low- and lower middle-income countries. For lower secondary education, access is still much

---

1 The 130 million girls out of school include 32 million out of primary, 29 million out of lower secondary, and 69 million out of upper secondary.
more limited for both boys and girls, with only 50 percent of girls and boys finishing lower secondary education in low and middle-income countries. Even at these low overall levels, many countries see girls at a disadvantage relative to boys (Figure 1B).

Even for those girls who are in school, many may not acquire the foundational cognitive skills to help them thrive in their future lives. Across 51 low- and middle-income countries, only 50 percent of young adult women who had completed grade 6 could read a simple sentence (Sandefur, Oye, and Pritchett 2016). The persistence in attainment gaps and low achievement in education translates into girls exiting school with fewer than optimal skills, reinforcing subsequent gaps in productive opportunities. Part of the solution is to ensure that girls get into school and get the most out of their schooling.

At the same time, evidence on what works to improve the quality of education is accumulating at an unprecedented rate (Figure 2). In recent years, hundreds of impact evaluations in low- and middle-income countries have demonstrated the effectiveness – or lack thereof – of a range of interventions at improving education outcomes, for girls and boys (Evans and Popova 2016; J-PAL 2017). Reviews that examine the most effective ways to boost girls’ education tend to focus on interventions that target girls – for example, building girls’ latrines at schools and providing scholarships for girls (Filmer and Schady 2008; Garn et al. 2013) – potentially missing large educational benefits for girls from interventions that are not gender-specific.

Consider two interventions: Intervention 1 targets girls and is effective for girls but not for boys – for example, providing scholarships targeted to high-performing girls. Intervention 2 is a general education intervention, which is much more effective for girls than Intervention 1, but the effects are roughly equal for boys and girls – for example, training teachers in an innovative way. In this case the most effective intervention for improving girls’ education is the general education intervention (Figure 3). The fact that it also benefits boys does not take away from the fact that it is the most effective intervention for girls. However, the published evaluation of Intervention 2 would likely not emphasize girls’ education or gender effects, as there are no differential effects for girls. In fact, the evaluation may not even report gender impacts separately.

Now consider two concrete interventions that demonstrate the above hypothesis (Figure 3). A merit scholarship program for girls in Kenya increased girls’ learning by 0.19 standard deviations and had small, positive spillovers for boys (Kremer, Miguel, and Thornton 2009). Another intervention – in Bangladesh – encouraged parents to attend monthly meetings with teachers to go over their child’s school performance. That intervention improved learning outcomes by 0.25 standard deviations for both girls and boys, with no statistically significant difference between them (Islam 2019). The general intervention (parent-teacher meetings) was more effective at increasing girls’ learning than the girl-targeted merit scholarship program. But it would likely be omitted in any review of the most effective interventions for

---

2 In education systems with high levels of attainment, the education gap can flip to disadvantage boy students (Psaki, McCarthy, and Mensch 2018).
girls because the study does not specifically target girls; and since the differences by gender are small, they do not come up in the title, the abstract, or the introduction of the paper. A traditional review of the best investments for girls, searching for studies with keywords like “girls” and “gender,” could miss the best investments for girls.

In this paper, we report the results of an innovative systematic review identifying the most effective interventions to improve girls’ access to education and learning outcomes within an evidence base that includes both girl-targeted and general education interventions. We pose three research questions: (1) Are girl-targeted interventions more effective for girls’ outcomes than general interventions? (2) For general, non-targeted interventions, do impacts on girls tend to be larger? and (3) In absolute terms, what are the most effective interventions for girls?

To answer these questions, we collected and reexamined a large database of education studies with access or learning outcomes for students. We categorized the studies as either evaluating girl-targeted or non-targeted (i.e., general) interventions and identified all studies that reported gender-differentiated impacts. Only one in three studies of interventions not targeted to girls report disaggregated impacts by gender, so a first implication of this work is that in order to understand how best to improve girls' education, studies should consistently report impacts for girls. For those studies that did not report gender differentiated impacts, we contacted their authors asking them either to run the additional gender differentiated analysis or to share the data with us. We then standardized the effects of different programs to ensure comparability of effect sizes across studies. Ultimately, we synthesized the effects for girls from 177 studies. (The full list of studies is available in Appendix A.)

We find that general, non-targeted interventions perform similarly to girl-targeted interventions on average to increase both girls' access to school and their learning in school. General interventions tend to have similar effects for girls and for boys. (The evidence suggests that if anything, girls benefit more from general interventions, although the differences are not statistically significant.) In examining the most and least effective interventions for girls’ education, we find that girls’ access to school is more responsive to changes in costs, distance, and health conditions; while girls’ learning is more likely to be improved by structured pedagogy and interventions that help teachers to teach at the right level. Later in the paper, we discuss the implications for inequality between boys and girls and cost-effectiveness of programs.

General, non-targeted interventions may be more politically palatable for scaling up, and they offer a wider array of evaluated interventions, giving policy makers a richer menu of options among non-targeted interventions to improve girls’ education. In countries where boys also struggle to achieve quality education, general interventions can simultaneously improve girls’ learning while benefitting boys as well. None of this suggests that programs will not benefit from considering gender issues in their design. Furthermore, many of the interventions included in this review focus on primary education, and as girls reach adolescence, they may face more gender-specific constraints. But for achieving access and quality, especially in primary education, specifically targeting girls may not always be necessary to help those girls succeed. If policy makers want to help girls learn, they can make
schools better for all children. Furthermore, programs can be mindful of gender issues without being specifically targeted to girls.

2. Literature review: The missing evidence

Various researchers have undertaken systematic reviews of evidence on what works to improve girls’ education. For example, Sperling and Winthrop (2015) – building on Herz and Sperling (2004) – summarize eight recent reviews of education impact evaluations and provide a catalogue of information about practices and programs that improve girls’ access and learning. Unterhalter et al. (2014) conduct a systematic review of education programs that target girls. Tembon and Fort (2008) document effective polices in six low- and middle-income countries to promote gender equality. J-PAL (2017) compares impacts by gender of programs – all evaluated by randomized controlled trial – designed to increase school enrollment and attendance in 28 low- and middle-income countries. Other reviews focus on adolescent girl programs that may also impact education outcomes (Haberland, McCarthy, and Brady 2018; Botea et al. 2017). As discussed above, these reviews – with the exception of J-PAL (2017) – search for studies that highlight girls’ education in the title, abstract, introduction, or in the characterization of the interventions, and so they largely exclude non-targeted interventions. For example, a common approach is to use terms such as “girls” or “gender” in the search for relevant studies and then to identify lessons learned from those studies, as do Unterhalter et al. (2014) and Haberland, McCarthy, and Brady (2018).

In addition, all the reviews except J-PAL (2017) and Unterhalter et al. (2014) examine the evidence qualitatively, usually discussing study by study, and then infer conclusions. J-PAL (2017) focuses only on access outcomes and within that, only on randomized controlled trials. Unterhalter et al. (2014) employs a form of vote-counting.3 To the best of our knowledge, there have been no systematic reviews that compare effect sizes quantitatively for both access and learning outcomes for girls’ education. This systematic review complements the existing literature in two ways: first, it incorporates a wide range of general education interventions that have not been included in previous studies of what works to improve girls’ learning and access to education; second, it converts the point estimates of all included studies into standardized effect sizes to allow easier comparison across studies.

3. Method

We gathered a large collection of studies that report education outcomes, either access or learning. For each of the studies, we identified whether or not they separately report impacts for boys and girls. For studies that separately report impacts for boys and girls, we extract those data, standardize the estimates, and use them to compare the impacts for boys versus girls and across programs for girls. For studies that do not separately report, we contacted

---

3 Vote counting consists of categorizing the results of studies by their results: negative and statistically significant, negative and statistically insignificant, positive and statistically insignificant, and positive and significant. Evans and Popova (2016) includes a discussion of different types of systematic reviews and the respective advantages and disadvantages.
the authors and asked them either to share the data or to provide the separate estimates themselves. In this section, we report on each step in detail.

**Literature search**

We began with a comprehensive database of education impact evaluations compiled for Evans and Popova (2016) and subsequently updated it. The database consists of 495 studies that were cited in 10 recent systematic reviews of evidence on what works to improve learning and access in low- and middle-income countries. All the reviews were published or made publicly available between 2013 and 2015 and the studies included were conducted between 1980 and 2015. Another systematic review of interventions with a special focus on access outcomes came out in 2017 (J-PAL 2017); its references added four studies to the database.

To increase the coverage of studies that were published (either as working papers or peer-reviewed articles) after 2015, we conducted an additional literature search between October 2017 and January 2018. We searched Google Scholar and the websites of major institutions that conduct research related to low- and middle-income countries for working papers that were published between 2015 and 2017 containing the keywords “evidence”, “education”, “access”, “learning”, “enrollment”, “dropout”, “attendance”, or “score”. We applied the same search terms to a number of economics and education journals, listed in Appendix B. These two additional searches yielded 19 new studies. In total, we reviewed 518 papers.

**Inclusion criteria**

We included studies of education interventions (such as teacher professional development and providing textbooks), health interventions (such as providing deworming drugs and micronutrients), and safety net interventions (such as cash transfers). We only included studies that took place in preprimary, primary, and secondary schools in low- or middle-income countries, according to the World Bank definition (World Bank 2017). To be included studies had to be published – either as a working paper or a journal article – between 1980 and 2017 and had to report at least one of the following education outcomes: access outcomes (enrollment, dropout, or attendance) or learning outcomes (composite test score or any subject score). Non-academic skill development programs for adolescents were not included.

---

4 The 10 reviews are: Conn (2014), Glewwe et al. (2014), Kremer et al. (2014), Krishnaratne et al. (2013), McEwan (2015), Murnane and Ganimian (2014), Asim et al. (2015), Masino and Niño-Zarazúa (2015), Glewwe and Muralidharan (2015), and Snijders et al. (2015). Conn, Glewwe et al., McEwan, and Masino and Niño-Zarazúa only include studies with learning outcomes. The other reviews include studies with learning outcomes and studies with access outcomes. The database is available at [https://sites.google.com/site/davidkevans/database-of-education-studies](https://sites.google.com/site/davidkevans/database-of-education-studies).
We only included studies that used an experimental or quasi-experimental design. To be included, studies needed to have a valid counterfactual – in other words, a credible way of determining what would have happened in the absence of the program. The ways that studies could construct such a counterfactual included random assignment of treatment, difference-in-differences analysis, regression discontinuity, instrumental variables, and propensity score matching. At the same time, we restrict our analysis to studies where girls are included in the intervention group.

**Data collection**

Upon reviewing the 518 identified studies, 330 studies met the inclusion criteria. We further divided these studies into two groups: girl-targeted interventions and general interventions. Girl-targeted interventions include any intervention that is explicitly designed to boost education outcomes for girls specifically. For example, this includes programs that provide girls with cash or in-kind transfers, reduce tuition or other school costs for girls, offer (merit) scholarships to girls, build latrines for girls in schools, reduce travel distance to schools for girls by building village schools or providing transportation, provide female teachers, or implement girls’ empowerment curricula in schools. In general, if the program either specifically targets girls for benefits or explicitly states its objective as improving girls’ educational outcomes, we count it as “girl-targeted.” In our sample, we identified 20 studies designed to increase access or learning specifically for girls. The other 310 studies were general interventions.

General interventions refer to programs that are gender neutral in their design. Examples include programs that offer computer-assisted learning for all students, provide school meals for all students, and distribute free school uniforms or textbooks to all students. A general intervention may disproportionately benefit girl students, but it is not explicitly designed to do so, nor is it targeted specifically to girls.

To collect the impacts of interventions on girls for the 20 girl-targeted studies, we used the results on girls reported in the studies. For general interventions, the average effect reported in the study covered an average across boys and girls, so we verified which studies also reported effects separately. 105 studies reported heterogeneous intervention impacts by gender in their original papers, and we incorporated those results in our review. However, that left 205 studies that did not report gender differentiated impacts in their original papers. In order to enlarge our sample, we contacted the authors of these studies up to three times between January 2018 and July 2018, requesting that they either provide additional estimates of intervention effects by gender or share the data of their studies with us to perform the analysis on their behalf. Authors were given at least three months to reply to us with either new estimates or their data if they were interested.

---

5 There are three general intervention studies that contain a girl-targeted intervention arm, but for the purpose of counting, but because the bulk of the benefits do not target girls, we include them in the general intervention group.
Of the 205 studies, we received replies from the authors of 104 studies. Among them, the authors of 32 studies indicated that the data were no longer available or that gender data were not collected. Another 72 sets of authors expressed their willingness to run the additional analysis (50 papers) or share their data with us (22 papers). By the end of July 2018, we were able to obtain new estimates of effects by gender of 52 studies.

Figure 4 demonstrates our review process. Combining girl-targeted interventions, general interventions that report impacts on girls and the new estimates we collected from authors, the final sample of this review consists of 177 studies evaluating 270 total interventions. Among those studies, 85 measured access outcomes such as enrollment, attendance, or dropout; and 114 measured learning outcomes including a composite test score, math score, or language score. Table 1A shows the descriptive statistics of our sample. Table 1B shows characteristics of the evaluated interventions. For access, more than half of general interventions targeted children of primary school age, whereas about one-fifth of girl-targeted interventions targeted girls of primary school age, and one-third focused on girls of secondary school age. The majority of general and girl-targeted interventions took place in both urban and rural settings, and just over half of the interventions had just one component. For learning, the majority of general interventions were implemented at the primary level and in both urban and rural areas. Seventy percent of girl-targeted learning interventions took place in rural areas. Sixty percent of all learning interventions had multiple components.

**Coding of effect sizes**

In this paper, our unit of analysis is the estimated impact of an intervention, where a group that received an intervention is compared to another group that did not receive the intervention. For studies with multiple treatment arms, we coded the impact of each treatment arm separately (as its own intervention) and recorded the education outcomes corresponding to that intervention. For example, Berlinski et al. (2016) tested the effects of four interventions or treatment arms: (1) an active learning approach to the teaching of math, (2) an active learning approach plus an interactive white board, (3) an active learning approach plus a computer lab, and (4) an active learning approach plus one computer per student. We coded these four experiments as four separate interventions. Furthermore, if studies reported multiple estimates for a given intervention, we coded all of those estimates separately rather than creating a composite variable.

Since studies in our sample collected different outcomes using different measures, in order to compare the effectiveness of the interventions on the same scale, individual point estimates need to be standardized. In this paper, we used Cohen’s $d$ to standardize effect sizes, following McEwan (2015) and Conn (2017). Cohen’s $d$ can be estimated using the raw mean difference between a treatment group ($\bar{Y}_T$) and a control group ($\bar{Y}_C$) as well as the
pooled standard deviation for the treatment and control groups combined ($S_{pooled}$) (see Equation 1).

$$d = \frac{\bar{Y}_T - \bar{Y}_C}{S_{pooled}}$$

(Equation 1)

In cases where the pooled standard deviation was not directly reported in the study, we calculated it using Equation 2 from Borenstein et al. (2009):

$$S_{pooled} = \sqrt{\frac{(n_T-1)S_T^2 + (n_C-1)S_C^2}{n_T + n_C - 2}}$$

(Equation 2)

where $n_T$ and $n_C$ are the sample sizes in the treatment and control groups, and $S_T$ and $S_C$ are the standard deviations in each group.

**Qualitative variables**

We collected a set of additional variables to better characterize the most effective interventions for girls. The variables included country, region, implementation agency, location (rural or urban), intervention level (village, school, household, or individual), duration of intervention (single contact or repeated contact), number of intervention components (single or multiple), the level of education at which the intervention was implemented (preprimary, primary, or secondary), student age, major program components (such as reducing school costs, a health intervention, additional teaching and learning materials, or school grants), the presence of components identified by program implementers as “girl friendly,” cost data (if any), quality of the outcome data (e.g., administrative data, self-reported data, national tests, international tests, program designed tests). For each study, we also coded its publication type and evaluation method.

**4. Results**

**Are girl-targeted interventions the most effective for girls?**

In terms of increasing girls’ participation in school, girl-targeted interventions and general interventions perform similarly on average, although there are some girl-targeted interventions that outstrip general interventions. Figure 5 demonstrates the distributions of effect sizes for girls from both general interventions and girl-targeted interventions. The median effect size for these two categories is very similar, increasing girls’ enrollment or attendance by 0.07-0.09 standard deviations. The effect sizes of less effective interventions – at the 10th and 25th percentiles – are also similar. However, the girl-targeted interventions at the 90th percentile have effect sizes that are 0.10 standard deviations larger than those of general interventions. That said, there are also general interventions with large effect sizes.

---

*In this paper, wherever applicable, we collected the mean difference with controls for observable variables.*
The effect size of the most effective general intervention (Conditional cash transfer in South Africa, Eyal, Woolard, and Burns (2014)) – 1.66 standard deviations – is comparable in size to that of the most effective girl-targeted intervention (Conditional cash transfer to girls in Malawi, Baird et al. (2016)), at 1.54 standard deviations. None of these differences are statistically significant at standard levels (Table 2).

There are also far more general interventions than specifically girl-targeted interventions. As we see in Table 1, the number of general interventions is more than three times that of girl-targeted interventions. This means that in each of the effect size bins, general interventions provide a larger menu for tested options (Figure 6). Even among the most effective interventions, there are almost as many general interventions with large effect sizes (greater than 0.4 standard deviations) because so many more general interventions have been tested. Therefore, general interventions constitute an important source of ways to improve girls’ access to education.

For learning, on average, girl-targeted and general interventions have comparable impacts on girls (Figure 7). The median interventions increase learning by 0.12 and 0.13 standard deviations. The top programs (90th percentile) of general interventions have slightly bigger measured effect sizes (0.52 standard deviations) than those of girl-targeted interventions (0.43 standard deviations), although the differences are not statistically significant.

However, as with the access studies, the difference in the number of general interventions and girl-targeted interventions is significant (Figure 8). This is even more the case in learning outcomes: there are 178 general learning interventions (from 106 studies) compared to only 14 girl-targeted learning interventions (Table 1A). With just 14 girl-targeted interventions (from 11 studies), the distribution of effect sizes might be affected by outliers: in fact, the large effect size of the top girl-targeted intervention (at the 90th percentile) is purely driven by a school construction intervention (Kazianga et al. 2013), which was designed specifically to increase girls’ access to schools.7 When schools were built in villages in rural Burkina Faso, learning outcomes for girls dramatically improved. But taking out this intervention, the effect size of girl-targeted interventions at the 90th percentile drops to 0.1 standard deviations. Alternatively, if one drops the largest two interventions from the general interventions, the effect size changes hardly at all.

These findings have two potential implications. The first is that while general and girl-targeted interventions perform similarly on average, there are more proven general interventions that deliver high impacts for girls’ learning than there are girl-targeted interventions. As a result, policy makers have more options to draw from among the general interventions. The second is that insofar as governments and other actors are experimenting with innovative girl-targeted interventions, there may be value in evaluating these to build the evidence base.

---

7 The Burkina Faso program, evaluated in (Kazianga et al. 2013), includes girls explicitly in the name of the program: the Burkinabe Response to Improve Girls’ Chances to Succeed.
For general interventions, do impacts on girls tend to be larger?

Previous research shows that the demand for girls’ schooling tends to be more responsive than that for boys’ to gender neutral education policies (Glick 2008; J-PAL 2017). However, we do not detect much difference between girls and boys in responding to interventions to increase access at any point in the distribution – the least effective programs, the median programs, and the most effective programs all have similar impacts in access for boys and girls (Figure 9). If anything, general interventions seem to be slightly more effective for girls than for boys, confirming previous work.8

Once in school, the impacts of general interventions on learning outcomes are also quite similar for girls and boys (Figure 10). Median interventions increased learning for girls and boys by 0.1 standard deviation for both, with slightly larger impacts for girls (Figure 10). At the top of the distribution (90th percentile), the effect for boys is 0.51 standard deviation, comparable to that for girls, at 0.52 standard deviation. None of these differences is statistically significant at standard levels (Table 2). These are sizeable effects which are the equivalent of more than 2.5 years of additional schooling in the business-as-usual setting in many low- and middle-income countries (Evans and Yuan 2019).

What are the most effective interventions for girls?

To summarize the most effective interventions for girls, in this section we present the 10 access and learning interventions with the largest effect sizes and seek to understand their attributes. We contrast these with the 10 least effective interventions in terms of access and learning outcomes. An alternative approach would be to carry out a formal meta-analysis: As the results demonstrate, there is a great deal of variation within categories of interventions (such as cash transfers), such that taking the average effect of a category is unlikely to yield meaningful insights.9 Because we code each estimate of the impact of an intervention separately, the fact that one estimate appears in the 10 most effective or least effective does not mean that all estimates of the impact of that intervention are among the most or least effective.

Access

The 10 studies that report the largest impacts in improving access to education for girls report greatly improved girls’ participation in school, with an average effect size of 0.74 standard deviation (Table 3). Three of the 10 are girl-targeted interventions, including cash transfers to girls who had previously dropped out of school – conditional on school

---

8 We also examine whether impacts on girls are larger in places with low levels of initial performance, using the harmonized learning indicators from the World Bank’s Human Capital Index (World Bank 2018a) as well as various access indicators from the World Development Indicators. We find no relationship.

9 Analysis of previous meta-analyses of education interventions suggests that high heterogeneity within categories limits the predictive power of meta-analysis in education (Masset 2019).
attendance in Malawi (Baird et al. 2016), improving school water and sanitation systems in Kenya (Garn et al. 2013), and providing private school subsidies for girls in Pakistan (Kim, Alderman, and Orazem 1999). Six of the general interventions are similarly related to offering cash for education in different countries (Eyal, Woolard, and Burns 2014; Maluccio, Murphy, and Regalia 2010; Edmonds and Shrestha 2014; Benhassine et al. 2015; Duflo, Dupas, and Kremer 2017), building village schools in Afghanistan (Burde and Linden 2013) and another intervention is focused on malaria prevention in The Gambia (Jukes et al. 2006). Altogether, six of the ten involve cash transfers, and one more – subsidies in Pakistan – similarly also involves reducing the cost of schooling.

These top interventions demonstrate that reducing the cost of schooling is likely the single most effective way to bring girls into school. Most of these are conditional cash transfers, although fewer unconditional transfers have been tested. In addition, reducing indirect costs – such as the commuting distance to school for girls by building village schools – has been effective in increasing access. Note, however, that one unconditional cash transfer – without a schooling condition – is among the less effective interventions (Table 4) (Baird, McIntosh, and Özler 2011). Improving health conditions through either better sanitation facilities or controlling malaria tends to attract more girls to school as well.

There are concerns about the effectiveness of conditional cash transfer programs if only considering the most effective interventions. One of them is that the popularity of conditional cash transfers has led to an emergence of impact evaluations in this field, which might lead this class of interventions to be overrepresented in the evidence base. Cash transfer interventions could be among both the most effective and the least effective interventions. To test this, we summarize the bottom 10 interventions to increase access for girls in Table 4. There are three transfer programs – conditional cash or in-kind – that were particularly ineffective in bringing girls into school, such as those in Burkina Faso (Kazianga et al. 2013), the Philippines (Chaudhury, Friedman, and Onishi 2013) and Uruguay (Amarante, Ferrando, and Vigorito 2013), but transfer programs represent far more of the most effective than the least effective programs. There is more variation in the least effective programs, ranging from providing school meals to targeted savings accounts for education. Interestingly, we see that within the same study (Garn et al. 2013), while promoting hygiene, improving water treatment, improved sanitation, and safe water storage in Kenyan primary schools is one of the best ways to increase girls’ enrollment, promoting hygiene and improving water storage alone actually reduced enrollment for girls. Although it is likely that girls are more responsive to sanitation conditions, different environments face different challenges: in Nepal providing sanitary products did not increase girls’ school attendance, likely in part because very few girls reported missing school due to a lack of sanitary products (Oster and Thornton 2011).

**Learning**

For learning, the average effect size of the top interventions for girls is 0.96 standard deviation (Table 5). Compared to access interventions, there is more variation in the design of learning interventions. First, only two in 10 studies are girl-targeted interventions. One of the two girl-targeted interventions is a public private partnership initiative in schools in
Pakistan providing a gender differentiated subsidy that increased girls’ test score by 0.77 standard deviation. The other intervention arm in the same initiative provided a gender neutral subsidy and also yielded sizeable effects, albeit smaller than the gender-differentiated one (Barrera-Osorio et al. 2017). The other top-ten girl-targeted intervention is the Afghan village school program for girls that delivered significant impacts on girls’ access and learning outcomes (Burde and Linden 2013). A general (non-targeted) community school program in Honduras greatly improved girls’ math score (Di Gropello and Marshall 2011).

Several of the most effective general interventions for girls among the top 10 involve structured pedagogy in early grades, or providing teachers with clear guidance on teaching or even scripted lesson plans. These interventions have been shown to be highly effective in several Sub-Saharan African countries including South Africa, Liberia and Kenya (Piper 2009; Piper, Zuilkowski, and Ong’ele 2016; Piper and Mugenda 2014; Piper and Medina 2010a). Another category of interventions that work well for girls (and boys) are those that help teachers to teach children at their current level of learning (e.g., teaching at the right level), either through diagnostic feedback or software as reported in Banerjee et al. (2016) and Imbrogno (2014).

On the other hand, the least effective programs for girls’ learning are all general interventions (Table 6). Various interventions actually had negative impacts on learning for girls compared to “schooling-as-usual,” but often, those same programs did not work for boys either. For example, technology interventions – whether substituting teachers with computers or providing students with laptops – did not help improve learning (Linden 2008; Sharma 2014). Although there are teacher professional development programs that work to improve student learning (Popova et al. 2018), our findings demonstrate that introducing new pedagogical methods through a short teacher training program is less likely to be effective to improve girls’ learning; and this is true no matter which education level the intervention targets (Berlinski and Busso 2017; Yoshikawa et al. 2015). In addition, school accountability interventions such as distributing school report cards to students and parents were not effective for girls, either in Sri Lanka or in Liberia (Aturupane et al. 2014; Piper and Medina 2010b).

5. Discussion

**Inequality**

Up until this point, this paper has focused on identifying the interventions that deliver the highest absolute learning gains for girls. An alternative approach would be to identify those programs that benefit girls most relative to boys. In other words, this approach would focus on closing inequalities (or increasing them, in contexts where girls are ahead in school) rather than merely improving girls’ access and learning without regard to boys’ performance. Figure 11 shows the gains in access and learning for boys and for girls. The programs with the most unequal impacts – both favoring girls and favoring boys – are general interventions. Almost all of the girl-targeted interventions for which we have data on both girls and boys have similar results for both genders, with slightly better results for girls. 11 in 20 girl-targeted
studies do not report outcomes for boys. If one’s objective were purely inequality reduction, then cash transfers in South Africa had dramatically larger access impacts on girls than on boys, despite not being gender-targeted (Eyal, Woolard, and Burns 2014). A mother tongue learning instruction in Kenya in the Lubukusu language had no discernible impact on boys’ learning but a sizeable impact for girls (Piper, Zuilkowski, and Ong’ele 2016). However, there are no clear patterns as to which classes of interventions are inequality enhancing versus inequality reducing. For this inequality analysis, we drop girl-targeted studies that do not report impacts for boys. If one were to assume that those girl-targeted studies that did not report outcomes for girls had zero impact on boys, then girl-targeted interventions would decrease inequality more than general interventions.

Costs

While we standardized effect sizes across interventions in this review, incorporating cost data would enhance the analysis, as the most effective programs may not be the most cost-effective and therefore not easy to scale up. However, despite a strong demand for cost data, few studies report them. McEwan (2015) in his review stated that 56 percent of studies reported no cost details, and most of the rest reported minimal information. We encountered similar problems when we tried to collect cost data. In addition, even when cost data are reported, they are often not comparable due to different accounting methods. Taking an early childhood development program in rural Mexico as an example, the cost per child estimated by World Bank researchers was $76 (Cárdenas, Evans, and Holland 2015), but when evaluated by another group of researchers at Brookings, the cost per child almost doubled to $174-$202 (Gustafsson-Wright, Boggild-Jones, and Gardiner 2017). Ideally, a separate initiative would collect cost data following a standard set of guidelines such as those laid out in Dhaliwal et al. (2013).

This review finds that general interventions are often comparable in impact to girl-targeted interventions in improving access to school and learning once at school. But if a policy maker’s primary concern is improving girls’ education, then perhaps investing in girl-targeted interventions would allow similar gains at much lower cost – i.e., just paying for the girls rather than girls and boys. This argument plays out differently for access versus learning interventions. For access interventions such as cash transfers, the cost could indeed be potentially reduced by targeting only girls. Indeed, several of the most effective general interventions were cash transfer programs that happened to not target girls specifically. One could imagine replacing those programs with girl-only cash transfer programs and potentially achieving similar gains. For learning interventions, such as structured pedagogy interventions, many are introduced at the level of the school, so that in mixed-gender schools, there is no clear cost gain to trying to limit the impact to girls only.

Program attributes

We gathered data on a number of program attributes with the aim to provide more information on the most effective programs. For example, the average program size of the most effective access interventions is 262 students and for learning interventions, is 556
students. With the exception of the cash transfer programs, all others among the ten most and least effective programs are pilot programs. This is a result of the fact that most interventions that are carefully evaluated tend to be pilots. Therefore, we cannot from this sample infer whether or not pilot programs are more effective than those that have been implemented at scale. Another attribute that we examined was the level of education that the top programs targeted. In terms of access interventions, 7 of the top 10 interventions targeted school-aged children in general, often between age 6 up to age 16 – working through the household rather than the school, trying to get out-of-school children into school. For learning interventions, 9 in 10 focused on the primary level, and half of them were designed to improve learning in grades 1-3. There is great interest in programs for adolescent girls, but many of those programs focus on building life skills and increasing earning capacity directly (see, for example, (Adoho et al. 2014; Bandiera et al. 2018; Bandiera et al. 2019) rather than keeping girls in school and increasing their learning ability. In many low- and middle-income countries, children and youth can still significantly improve their literacy and numeracy all through primary and secondary school (Evans and Yuan 2019), and so there will be great value in continuing to evaluate programs and increase learning and access for adolescents. We also examined if authors included any gender component in their interventions. We found that besides girl-targeted programs, only 1 general intervention in the top 20 had a girl-friendly component, which was to provide gender-differentiated school subsidies (Barrera-Osorio et al. 2017).

What has been studied

A key limitation of this work is that it only surveys those interventions that have been evaluated. One can imagine a wide array of girl-targeted interventions that could still be tried or that have been tried but not yet rigorously evaluated. In the context of strict budget constraints, having clear data on the best investments among those interventions that have been evaluated can be useful, and it can help governments and other education stakeholders to avoid investing in programs that have proven ineffective. However, it should not stop policy makers from continuing to innovate and test new programs that relax constraints on girls’ access and learning.

6. Conclusion

Previous reviews of what works to improve girls’ education tend to focus on girl-targeted interventions. That approach omits key evidence of the impact of general education interventions on girls. This review innovatively brings together a large evidence base of general interventions that report effects for girls. Based on 177 studies from 54 countries, this review finds that girls’ access to school is more responsive to changes in costs, distance, and health conditions; while girls’ learning is more likely to be improved by structured pedagogy and interventions that help teachers to teach at the right level.

While this review focuses on girls’ education, the global learning crisis impoverishes both girls and boys (World Bank 2018b). Our findings demonstrate that gender neutral interventions hold great promise for girls’ learning as well as for boys. Considering the
limited resources that education systems in most low- and middle-income countries possess, the most practical approach to help girls learn may be to make schools better for all children. Such an approach may also be more politically palatable to voters than programs that restrict their benefits to girls.

Finally, attending school and acquiring learning are not the finish line for girls’ education. The ultimate objective is that girls can empower themselves through education and achieve their life aspirations. To this point, very few evaluations have included either long-term follow-ups or these broader measures of well-being. But gaining literacy and numeracy are the foundation for positive longer term outcomes.

References


Gustafsson-Wright, Emily, Izzy Boggild-Jones, and Sophie Gardiner. 2017. 'SECT: The standardized early childhood development costing tool'.


ILO. 2017. 'World Employment and Social Outlook: Trends for Women 2017'.

Imbrogno, Jason. 2014. 'Essays on the Economics of Education'.

International Initiative for Impact Evaluation (3ie). 2016. 'Impact Evaluation Repository.'


———. 2010b. 'Early Grade Reading Assessment (EGRA) Plus: Liberia. Program evaluation report'.

Piper, Benjamin, Stephanie S Zuilkowski, and Salome Ong'ele. 2016. 'Implementing Mother Tongue Instruction in the Real World: Results from a Medium-Scale Randomized Controlled Trial in Kenya', Comparative Education Review, 60: 776-807.

Popova, Anna, David K Evans, Mary E. Breeding, and Violeta Arancibia. 2018. 'Global Landscape of Teacher Professional Development Programs: The Gap between Evidence and Practice'.


Sandefur, Justin, Mari Oye, and Lant Pritchett. 2016. 'Girls' Schooling is Good, Girls' Schooling with Learning is Better', Background paper for the International Commission on Financing Global Education Opportunity.


UNESCO. 2016 'Leaving no one behind: How far on the way to universal primary and secondary education?', Policy Paper 27/Fact Sheet 37.

World Bank. 2017. 'World Development Indicators'.
Bank: Washington, DC).
Yoshikawa, Hirokazu, Diana Leyva, Catherine E Snow, Ernesto Treviño, M Barata,
Christina Weiland, Celia J Gomez, Lorenzo Moreno, Andrea Rolla, and Nikhit D'Sa.
2015. 'Experimental impacts of a teacher professional development program in Chile on
preschool classroom quality and child outcomes', Developmental psychology, 51: 309.

**Figures and tables**

**Figure 1A: Gender gap in primary school completion**

Note: Any gap within three percent is considered as gender parity.
Figure 1B: Gender gap in lower secondary school completion

Note: Any gap within three percent is considered as gender parity.

Figure 2: Evidence on what works in education has mushroomed

Access

Learning

Source: Authors’ adaptation using International Initiative for Impact Evaluation (3ie) (2016) and Evans and Popova (2016).
Figure 3: Interventions not targeted specifically to girls may in fact deliver bigger gains to girls

A thought experiment

<table>
<thead>
<tr>
<th>Increased learning (standard deviations)</th>
<th>Girls</th>
<th>Boys</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girl-targeted intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General education intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Two real-life examples

<table>
<thead>
<tr>
<th>Increased learning (standard deviations)</th>
<th>Girls</th>
<th>Boys</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girl-targeted intervention: Incentives to Learn, Kenya (Kremer, Miguel and Thornton 2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General education intervention: Parent-teacher meetings and student outcomes, Bangladesh (Islam, 2016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Impact Evaluations of Educational Interventions from a recent database (Evans and Popova 2016 and 3ie 2017 and miscellany)

[330 studies]

Without gender-differentiated impacts

[205 studies]

Authors provide new estimates (43 studies) OR we run the regression using shared data (9 studies)

[52 studies]

General interventions with gender-differentiated impacts

[105 studies]

Girl-targeted interventions

[20 studies]

Extract point estimates and calculate effect sizes for 270 interventions from 177 studies

Figure 5: Effect sizes of access outcomes for girls

<table>
<thead>
<tr>
<th></th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.07</td>
<td>0.15</td>
<td>0.27</td>
<td>[-0.2, 1.66]</td>
</tr>
<tr>
<td>Girl-targeted</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.09</td>
<td>0.23</td>
<td>0.37</td>
<td>[-0.18, 1.54]</td>
</tr>
</tbody>
</table>
Figure 6: Number of access outcomes by effect size

General interventions

![Diagram showing the number of access outcomes by effect size for general interventions.]

Girl-targeted interventions

![Diagram showing the number of access outcomes by effect size for girl-targeted interventions.]

Figure 7: Effect sizes of learning outcomes for girls

<table>
<thead>
<tr>
<th></th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.12</td>
<td>0.27</td>
<td>0.52</td>
<td>[-0.61, 2.56]</td>
</tr>
<tr>
<td>Girl-targeted</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.13</td>
<td>0.23</td>
<td>0.43</td>
<td>[-0.03, 0.45]</td>
</tr>
</tbody>
</table>

Range
Figure 8: Number of learning outcomes by effect size

General interventions

<table>
<thead>
<tr>
<th>Effect size</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0</td>
<td>79</td>
</tr>
<tr>
<td>0.0-0.1</td>
<td>124</td>
</tr>
<tr>
<td>0.1-0.2</td>
<td>75</td>
</tr>
<tr>
<td>0.2-0.4</td>
<td>95</td>
</tr>
<tr>
<td>&gt;0.4</td>
<td>56</td>
</tr>
</tbody>
</table>

Girl-targeted interventions

<table>
<thead>
<tr>
<th>Effect size</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0</td>
<td>3</td>
</tr>
<tr>
<td>0.0-0.1</td>
<td>6</td>
</tr>
<tr>
<td>0.1-0.2</td>
<td>7</td>
</tr>
<tr>
<td>0.2-0.4</td>
<td>4</td>
</tr>
<tr>
<td>&gt;0.4</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 9: Effect sizes of access outcomes for girls and boys (general interventions only)

<table>
<thead>
<tr>
<th>Girls</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4 standard deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.07</td>
<td>0.15</td>
</tr>
<tr>
<td>25th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Boys</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4 standard deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td>25th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90th</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Figure 10:** Effect sizes of learning outcomes for girls and boys (general interventions only)

<table>
<thead>
<tr>
<th></th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.12</td>
<td>0.27</td>
<td>0.52</td>
<td>[-0.61, 2.56]</td>
</tr>
<tr>
<td>Boys</td>
<td>-0.06</td>
<td>0.01</td>
<td>0.10</td>
<td>0.25</td>
<td>0.51</td>
<td>[-0.59, 1.98]</td>
</tr>
</tbody>
</table>

**Figure 11:** Relative effect sizes for boys versus girls

Note: Blue points represent effect sizes of general interventions, red dots represent effects sizes of girl-targeted interventions, and the green dash line represents same effect sizes for girls and boys. For access outcomes, 172 effect sizes of general interventions and 12 (out of 35) effect sizes of girl-targeted interventions are plotted; for learning outcomes, 420 effect sizes of general interventions and 7 (out of 23) effect sizes of girl-targeted interventions are plotted. The missing girl-targeted interventions are those that did not report outcomes for boys.
### Table 1A: Descriptive statistics of the overall sample

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of studies</th>
<th>Number of interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>General studies</td>
<td>157</td>
<td>239</td>
</tr>
<tr>
<td>General studies – Access</td>
<td>69</td>
<td>95</td>
</tr>
<tr>
<td>General studies – Learning</td>
<td>106</td>
<td>178</td>
</tr>
<tr>
<td>Girl-targeted studies</td>
<td>20</td>
<td>31</td>
</tr>
<tr>
<td>Girl-targeted studies – Access</td>
<td>18</td>
<td>26</td>
</tr>
<tr>
<td>Girl-targeted studies – Learning</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>177</td>
<td>270</td>
</tr>
</tbody>
</table>

Note: Access and learning do not sum to the total because multiple studies report both learning and access outcomes.
<table>
<thead>
<tr>
<th></th>
<th>All interventions</th>
<th>General interventions</th>
<th>Girl-targeted interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Access</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>56 (47%)</td>
<td>50 (53%)</td>
<td>6 (23%)</td>
</tr>
<tr>
<td>Secondary (lower and/or upper)</td>
<td>18 (15%)</td>
<td>9 (9%)</td>
<td>9 (35%)</td>
</tr>
<tr>
<td>Both primary and secondary</td>
<td>27 (23%)</td>
<td>20 (21%)</td>
<td>9 (35%)</td>
</tr>
<tr>
<td>Other</td>
<td>18 (15%)</td>
<td>16 (17%)</td>
<td>2 (8%)</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>13 (11%)</td>
<td>11 (12%)</td>
<td>2 (8%)</td>
</tr>
<tr>
<td>Rural</td>
<td>39 (32%)</td>
<td>31 (33%)</td>
<td>8 (31%)</td>
</tr>
<tr>
<td>Both</td>
<td>69 (57%)</td>
<td>53 (56%)</td>
<td>16 (62%)</td>
</tr>
<tr>
<td>Components</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>66 (55%)</td>
<td>51 (54%)</td>
<td>15 (58%)</td>
</tr>
<tr>
<td>Multiple</td>
<td>55 (45%)</td>
<td>44 (46%)</td>
<td>11 (42%)</td>
</tr>
<tr>
<td><strong>Learning</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>124 (65%)</td>
<td>118 (66%)</td>
<td>6 (43%)</td>
</tr>
<tr>
<td>Secondary</td>
<td>30 (16%)</td>
<td>30 (17%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Both primary and secondary</td>
<td>24 (13%)</td>
<td>16 (9%)</td>
<td>8 (57%)</td>
</tr>
<tr>
<td>Other</td>
<td>14 (7%)</td>
<td>14 (8%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>22 (11%)</td>
<td>22 (12%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Rural</td>
<td>74 (39%)</td>
<td>64 (36%)</td>
<td>10 (71%)</td>
</tr>
<tr>
<td>Both</td>
<td>96 (50%)</td>
<td>92 (52%)</td>
<td>4 (29%)</td>
</tr>
<tr>
<td>Components</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>80 (42%)</td>
<td>74 (42%)</td>
<td>6 (43%)</td>
</tr>
<tr>
<td>Multiple</td>
<td>112 (58%)</td>
<td>104 (58%)</td>
<td>8 (57%)</td>
</tr>
</tbody>
</table>
Table 2: Tests for statistically significant differences in distribution of intervention effect sizes

<table>
<thead>
<tr>
<th></th>
<th>Access outcomes</th>
<th>Learning outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Mean within 50&lt;sup&gt;th&lt;/sup&gt; – 75&lt;sup&gt;th&lt;/sup&gt; percentile</td>
</tr>
<tr>
<td><strong>Girl-targeted vs. General interventions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girl-targeted</td>
<td>0.048</td>
<td>0.037***</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.111***</td>
<td>0.106***</td>
</tr>
<tr>
<td><strong>Girls vs. Boys</strong></td>
<td>Mean</td>
<td>50&lt;sup&gt;th&lt;/sup&gt; – 75&lt;sup&gt;th&lt;/sup&gt; percentile</td>
</tr>
<tr>
<td>Girl</td>
<td>0.033*</td>
<td>0.031***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.086***</td>
<td>0.079***</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>212</td>
<td>52</td>
</tr>
<tr>
<td><strong>Learning outcomes</strong></td>
<td>Mean</td>
<td>50&lt;sup&gt;th&lt;/sup&gt; – 75&lt;sup&gt;th&lt;/sup&gt; percentile</td>
</tr>
<tr>
<td><strong>Girl-targeted vs. General interventions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girl-targeted</td>
<td>-0.027</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.184***</td>
<td>0.190***</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>452</td>
<td>117</td>
</tr>
<tr>
<td><strong>Girls vs. Boys in general interventions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>0.010</td>
<td>0.019***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.173***</td>
<td>0.171***</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>879</td>
<td>222</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

### p<0.01, ** p<0.05, * p<0.1
Table 3: The 10 most effective interventions to improve access to education for girls

<table>
<thead>
<tr>
<th>Program description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation design</th>
<th>Outcome</th>
<th>Effect size (SD)</th>
<th>Type of publication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditional cash transfer</td>
<td>South Africa</td>
<td>SSA</td>
<td>DID</td>
<td>Enrollment</td>
<td>1.657</td>
<td>Working paper</td>
<td>Eyal, Woolard, &amp; Burns 2014</td>
</tr>
<tr>
<td>Conditional cash transfer for dropped out girls</td>
<td>Malawi</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment, Yr 2</td>
<td>1.536</td>
<td>Working paper</td>
<td>Baird et al. 2016</td>
</tr>
<tr>
<td>Conditional cash transfer</td>
<td>Nicaragua</td>
<td>LAC</td>
<td>DID</td>
<td>Enrollment, Yr 1</td>
<td>0.883</td>
<td>Journal</td>
<td>Maluccio et al. 2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Enrollment, Yr 2</td>
<td>0.617</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hygiene promotion + water treatment + sanitation + water supply</td>
<td>Kenya</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.634</td>
<td>Journal</td>
<td>Garn et al. 2013</td>
</tr>
<tr>
<td>Free secondary education</td>
<td>Ghana</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.634</td>
<td>Working paper</td>
<td>Duflo et al. 2017</td>
</tr>
<tr>
<td>Malaria prevention</td>
<td>Gambia</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment – cohort w/o contamination</td>
<td>0.555</td>
<td>Journal</td>
<td>Jukes et al. 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Enrollment – cohort w/ minimal contamination</td>
<td>0.457</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional cash transfer: school stipend</td>
<td>Nepal</td>
<td>SA</td>
<td>RCT</td>
<td>Attendance rate</td>
<td>0.517</td>
<td>Journal</td>
<td>Edmonds and Shrestha 2014</td>
</tr>
<tr>
<td>Labelled cash transfer for education</td>
<td>Morocco</td>
<td>MENA</td>
<td>RCT</td>
<td>Dropped out by the end of year 2</td>
<td>0.486 (abs. value)</td>
<td>Journal</td>
<td>Benhassine et al. 2015</td>
</tr>
<tr>
<td>Village-based school</td>
<td>Afghanistan</td>
<td>SA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.478</td>
<td>Journal</td>
<td>Burde &amp; Linden 2013</td>
</tr>
<tr>
<td>Private school subsidies for girls</td>
<td>Pakistan</td>
<td>SA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>0.441</td>
<td>Journal</td>
<td>Kim et al. 1999</td>
</tr>
</tbody>
</table>

Note: Interventions marked with an asterisk are girl-targeted.
Table 4: The 10 least effective interventions to improve access to education for girls

<table>
<thead>
<tr>
<th>Program description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation design</th>
<th>Outcome</th>
<th>Effect size (SD)</th>
<th>Type of publication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>School canteen</td>
<td>Burkina Faso</td>
<td>SSA</td>
<td>RCT</td>
<td>Absenteeism</td>
<td>-0.200</td>
<td>Journal</td>
<td>Kazianga et al. 2012</td>
</tr>
<tr>
<td>*Conditional take-home rations for girls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional cash transfer to girls</td>
<td>Malawi</td>
<td>SSA</td>
<td>RCT</td>
<td>Attendance, Yr2</td>
<td>-0.152</td>
<td>Journal</td>
<td>Baird et al. 2011</td>
</tr>
<tr>
<td>Early financial commitment</td>
<td>China</td>
<td>EAP</td>
<td>RCT</td>
<td>Dropout</td>
<td>-0.140¹</td>
<td>Journal</td>
<td>Yi et al. 2015</td>
</tr>
<tr>
<td>Hygiene promotion + water treatment</td>
<td>Kenya</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment</td>
<td>-0.138</td>
<td>Journal</td>
<td>Garn et al. 2013</td>
</tr>
<tr>
<td>School meal</td>
<td>India</td>
<td>SA</td>
<td>DID</td>
<td>Enrollment</td>
<td>-0.120</td>
<td>Journal</td>
<td>Afridi 2011</td>
</tr>
<tr>
<td>Education cash saving account with parent outreach</td>
<td>Uganda</td>
<td>SSA</td>
<td>RCT</td>
<td>Enrollment Y1</td>
<td>-0.110</td>
<td>Working paper</td>
<td>Karlan and Leiden 2014</td>
</tr>
<tr>
<td>Education cash saving account without parent outreach</td>
<td></td>
<td></td>
<td></td>
<td>Enrolled in school 15-17 yrs</td>
<td>-0.098</td>
<td>Evaluation report</td>
<td>Chaudhury et al. 2013</td>
</tr>
<tr>
<td>Conditional cash transfer</td>
<td>Philippines</td>
<td>EAP</td>
<td>RCT</td>
<td>Attendance</td>
<td>-0.083</td>
<td>Journal</td>
<td>Oster and Thornton 2011</td>
</tr>
<tr>
<td>Sanitary products</td>
<td>Nepal</td>
<td>SA</td>
<td>RCT</td>
<td>Attendance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional cash transfer</td>
<td>Uruguay</td>
<td>LAC</td>
<td>DID</td>
<td>School attendance, 18-month</td>
<td>-0.056</td>
<td>Journal</td>
<td>Amarante et al. 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>follow up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unconditional Cash Transfer</td>
<td>South Africa</td>
<td>SSA</td>
<td>DID</td>
<td>Attendance</td>
<td>-0.043</td>
<td>Working Paper</td>
<td>Santana 2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Interventions marked with an asterisk are girl-targeted.

¹ Adjusted negative value for comparison.
Table 5: The 10 most effective interventions to improve learning for girls

<table>
<thead>
<tr>
<th></th>
<th>Program description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation design</th>
<th>Outcome</th>
<th>Effect size (SD)</th>
<th>Type of publication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Literacy intervention</td>
<td>South Africa</td>
<td>SSA</td>
<td>DID</td>
<td>Letter sounding fluency</td>
<td>2.563</td>
<td>Evaluation report</td>
<td>Piper 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Word naming fluency</td>
<td>1.840</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reading comprehension</td>
<td>1.757</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Oral reading fluency</td>
<td>1.658</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Mother tongue instruction</td>
<td>Kenya</td>
<td>SSA</td>
<td>RCT</td>
<td>Reading comprehension, Lubukusu, class 1</td>
<td>1.36</td>
<td>Journal</td>
<td>Piper et al. 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reading comprehension, Kikamba, class 2</td>
<td>1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>TaRL 10-day Camp</td>
<td>India</td>
<td>SA</td>
<td>RCT</td>
<td>Language</td>
<td>1.050</td>
<td>Working paper</td>
<td>Banerjee et al. 2016</td>
</tr>
<tr>
<td></td>
<td>TaRL 10-day Camp</td>
<td></td>
<td></td>
<td></td>
<td>Math</td>
<td>0.870</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TaRL 20-day Camp</td>
<td></td>
<td></td>
<td></td>
<td>Language</td>
<td>0.830</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TaRL 20-day Camp</td>
<td></td>
<td></td>
<td></td>
<td>Math</td>
<td>0.730</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Structured pedagogy</td>
<td>Liberia</td>
<td>SSA</td>
<td>DID</td>
<td>Listening comprehension</td>
<td>1.030</td>
<td>Evaluation report</td>
<td>Piper &amp; Korda 2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reading comprehension</td>
<td>0.830</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unfamiliar word fluency</td>
<td>0.780</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Letter-naming fluency</td>
<td>0.680</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Oral reading fluency</td>
<td>0.680</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Kenya</td>
<td>SSA</td>
<td>SSA</td>
<td>RCT</td>
<td>English letter sound</td>
<td>0.845</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>Country</td>
<td>Region</td>
<td>Study Design</td>
<td>Outcome</td>
<td>Score</td>
<td>Evaluation Source</td>
<td>Authors</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td>--------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>-------</td>
<td>-------------------------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>Primary literacy intervention</td>
<td></td>
<td></td>
<td></td>
<td>English segmenting</td>
<td>0.690</td>
<td>Evaluation report</td>
<td>Piper &amp; Mugenda 2014</td>
<td></td>
</tr>
<tr>
<td>*PPP gender subsidy</td>
<td>Pakistan</td>
<td>SA</td>
<td>RCT</td>
<td>Test score</td>
<td>0.770</td>
<td>Working paper</td>
<td>Barrera-Osorio et al. 2017</td>
<td></td>
</tr>
<tr>
<td>PPP subsidy pooled</td>
<td></td>
<td></td>
<td></td>
<td>Test score</td>
<td>0.661</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPP uniform subsidy</td>
<td></td>
<td></td>
<td></td>
<td>Test score</td>
<td>0.655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Village-based schools</td>
<td>Afghanistan</td>
<td>SA</td>
<td>RCT</td>
<td>Test score - 2nd semester</td>
<td>0.661</td>
<td>Journal</td>
<td>Burde &amp; Linden 2013</td>
<td></td>
</tr>
<tr>
<td>Math tutor software</td>
<td>Mexico</td>
<td>LAC</td>
<td>RCT</td>
<td>Test scores</td>
<td>0.660</td>
<td>Working paper</td>
<td>Imbrogno 2014</td>
<td></td>
</tr>
<tr>
<td>Community School Program</td>
<td>Honduras</td>
<td>LAC</td>
<td>IV</td>
<td>Math</td>
<td>0.630</td>
<td>Journal</td>
<td>Di Gropello &amp; Marshall 2011</td>
<td></td>
</tr>
<tr>
<td>Math tutor software</td>
<td>Chile</td>
<td>LAC</td>
<td>RCT</td>
<td>Test scores</td>
<td>0.611</td>
<td>Working paper</td>
<td>Imbrogno 2014</td>
<td></td>
</tr>
</tbody>
</table>

Note: Interventions marked with an asterisk are girl-targeted.

1 TaRL: Teaching at the Right Level
Table 6: The 10 least effective interventions to improve learning for girls

<table>
<thead>
<tr>
<th>Program description</th>
<th>Country</th>
<th>Region</th>
<th>Evaluation design</th>
<th>Outcome</th>
<th>Effect size (SD)</th>
<th>Type of publication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer assisted learning in school</td>
<td>India</td>
<td>SA</td>
<td>RCT</td>
<td>Math and English</td>
<td>-0.613</td>
<td>Working Paper</td>
<td>Linden 2008</td>
</tr>
<tr>
<td>School management</td>
<td>Madagascar</td>
<td>SSA</td>
<td>RCT</td>
<td>Test score (district level intervention)</td>
<td>-0.403</td>
<td>Journal</td>
<td>Glewwe and Maiga 2011</td>
</tr>
<tr>
<td>New curriculum + OLPC</td>
<td>Costa Rica</td>
<td>LAC</td>
<td>RCT</td>
<td>Math-geometry</td>
<td>-0.378</td>
<td>Journal</td>
<td>Berlinski and Busso 2017</td>
</tr>
<tr>
<td>New curriculum + computer lab</td>
<td></td>
<td></td>
<td></td>
<td>Math-geometry</td>
<td>-0.216</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New curriculum</td>
<td></td>
<td></td>
<td></td>
<td>Math-geometry</td>
<td>-0.142</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New curriculum + white board</td>
<td></td>
<td></td>
<td></td>
<td>Math-geometry</td>
<td>-0.136</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher training</td>
<td>Chile</td>
<td>LAC</td>
<td>RCT</td>
<td>Vocabulary</td>
<td>-0.305</td>
<td>Journal</td>
<td>Yoshikawa et al. 2015</td>
</tr>
<tr>
<td>One Laptop per Child</td>
<td>Nepal</td>
<td>SA</td>
<td>DID</td>
<td>English</td>
<td>-0.244</td>
<td>Working paper</td>
<td>Sharma 2014</td>
</tr>
<tr>
<td>Mobile school librarian</td>
<td>India</td>
<td>SA</td>
<td>RCT</td>
<td>Language</td>
<td>-0.232</td>
<td>Working paper</td>
<td>Borkum and Linden 2013</td>
</tr>
<tr>
<td>Preschool voucher</td>
<td>China</td>
<td>EAP</td>
<td>RCT</td>
<td>Test score - school readiness</td>
<td>-0.223</td>
<td>Journal</td>
<td>Wong et al. 2013</td>
</tr>
<tr>
<td>School report card</td>
<td>Sri Lanka</td>
<td>SA</td>
<td>DID</td>
<td>Science 8th grade</td>
<td>-0.221</td>
<td>Journal</td>
<td>Aturupane et al. 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Math 8th grade</td>
<td>-0.215</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attendance reward</td>
<td>India</td>
<td>SA</td>
<td>RCT</td>
<td>Test score</td>
<td>-0.207</td>
<td>Working paper</td>
<td>Visaria et al. 2016</td>
</tr>
<tr>
<td>Teacher training</td>
<td>Malawi</td>
<td>SSA</td>
<td>RCT</td>
<td>Early Grade Math, 36-month follow-up</td>
<td>-0.124</td>
<td>Working paper</td>
<td>Ozler et al. 2016</td>
</tr>
</tbody>
</table>
Appendix A. Studies included

This appendix lists the 177 studies included in the review. The 20 studies that examine girl-targeted programs are marked with an asterisk.


Battaglia, Marianna, & Lebedinski, Lara. (2014). *The curse of low aspirations: Remedial education and perceived returns to education of roma people.* Retrieved from


Lakshminarayana, Rashmi, Eble, Alex, Bhakta, Preetha, Frost, Chris, Boone, Peter, Elbourne, Diana, & Mann, Vera. (2013). The support to rural India's public education system (STRIPES) trial: a cluster randomised controlled trial of supplementary teaching, learning material and material support. *PloS one, 8*(7), e65775.


Appendix B. Journals and websites searched

Journals searched:

- *American Economic Review,*
- *Comparative Education Review,*
- *Economic Development and Cultural Change,*
- *Journal of Development Economics,*
- *Journal of Development Effectiveness,*
- *International Journal of Educational Development,*
- *Journal of Human Resources,*
- *Quarterly Journal of Economics,*
- *The Economics of Education Review,*
- *World Bank Economic Review,*
- *Journal of Public Economics,*
- *Journal of Human Capital.*

Websites searched:

- 3ie impact evaluation database,
- RISE Programme Conference 2017,
- CASE 2017 Conference,
- the Abdul Latif Jameel Poverty Action Lab at MIT;
- Innovations for Poverty Action;
- the Inter-American Development Bank;
- the National Bureau of Economic Research;
- RTI International;
- the Rural Education Action Program at Stanford University;
- the World Bank;
- the IZA Institute of Labor Economics.