Better learning for a sustainable future: A systematic review of education programmes

Birte Snilstveit† (Evaluation Specialist, 3ie), Martina Vojtkova (Evaluation Specialist, 3ie), Jennifer Stevenson (Research Associate, 3ie), Daniel Phillips (Research Associate, 3ie), Emma Gallagher (Research Associate, 3ie), Tanja Schmidt (Research Assistant, 3ie), Hannah Jobse (Research Assistant, 3ie), Maisie Geelen (Research Assistant, 3ie), Maria Grazia Pastorello (Research Assistant, 3ie), John Eyers (Information Specialist, 3ie), Philip Davies (Deputy Director, 3ie)

1. Introduction

Substantial progress has been made in improving access to education in low- and middle-income countries (L&MICs), but this has been uneven and challenges remain. In 2010, 61 million children of primary school age were still out of school, more than half of them (33 million) in Sub-Saharan Africa (UNESCO, 2012). Moreover, while there has been progress in reducing the number of girls excluded from education, from 58 per cent in 1999 to 53 percent in 2010 (UN, 2012), girls are still more likely than boys to miss out on schooling.

Moreover, improved access to education has failed to translate into learning in many countries (Prichett, 2013). According to the Education for All Global Monitoring report (UNESCO, 2013) around 250 million children in L&MICs cannot read, write, or do basic maths problems. This number includes over 130 million children who are actually enrolled in primary school and yet have not acquired these basic skills, leading some commentators to suggest there is a global learning crisis (Robinson, 2011). Thus, researchers and decision makers are increasingly focusing their attention on efforts to improve learning for all (Prichett, 2013; Robinson, 2011).

A range of different programmes are implemented to address the challenges associated with ensuring all children in L&MICs have access to schooling, and that they gain sufficient skills and knowledge to realise the benefits a good education can bring. As decision-makers are focusing their attention on improving learning for a sustainable future there is a need to identify promising approaches to ensuring high quality education for all.

To help inform decisions about how to use limited resources to achieve this goal, we conducted a systematic review of the evidence on the relative effects of education interventions in improving education access and learning outcomes for primary and secondary school children in L&MICs. In this paper we summarise the main findings from this study. The next section briefly summarise the objectives, methods and scope of the systematic review. The third section provides the results of the review, followed by a brief discussion and conclusion.

2. Objectives, methods and scope

The primary objective of this study was to identify, assess and synthesise evidence on the effects of different education interventions on children’s access to education and learning in

---

† Corresponding author: bsnilstveit@3ieimpact.org, International Initiative for Impact Evaluation (3ie), London International Development Centre, 36 Gordon Square, WC1H 0PD, London

Due to the broad scope of the study and word limit for this paper we are only able to provide a summary of main findings here. The full technical report from this study will be published in late 2015.
We conducted a systematic review following published guidelines (Hammerstrøm et al., 2010; Higgins & Green, 2011; The Campbell Collaboration, 2015) and synthesised studies using meta-analysis when feasible.

We included studies of primary and secondary school children in mainstream education in LMICs published between 1990 – January 2014. To be included studies had to use an experimental or quasi-experimental study design\(^4\) and measure school participation (enrolment, attendance, drop-out, completion), learning (cognitive skills, maths, language arts and composite score), teacher attendance or teacher performance.

We aimed to include a comprehensive range of commonly implemented education interventions designed to address one or more barriers to school participation and learning. Figure 1 provides an overview of included interventions classified according to the settings in which they take place and outline the main pathways through which they may improve education outcomes.

For example, a child-centred intervention such as school feeding provides an incentive for parents to send their children to school, but also aims to provide students with the nutrition they need to learn. Better nourished children are less likely to miss school due to illness, and better attendance can have the knock-on effect of improving learning outcomes (Kristjansson et al., 2009). Providing information, either to children or parents, is intended to emphasise the long-term benefits that schooling provides and, in so doing, increase demand for both enrolment and attendance (Krishnaratne & White, 2013). School-based management interventions are intended to improve the efficiency of school administration and leadership by facilitating innovation and allowing parent power to drive up the quality of schooling (Banerjee et al., 2008). Increasing quality of schooling and improved learning outcomes may also have an important impact on enrolment and attendance - and vice-versa. Improving the learning environment should have a direct impact on learning outcomes but may also have the indirect effect of pushing up demand by increasing the perceived benefits of schooling. Greater enrolment and attendance may change the student-teacher ratio, or lead to greater competition for limited resources or the inclusion of more children with a lower educational baseline.

The full details of the review objectives, methods and scope is available elsewhere (Snilstveit et al., 2015).

\(^3\) The review also assesses the effect of interventions on different sub-groups, as well as addressing a secondary objective on intervention implementation based on a review of qualitative studies and process evaluation. The level of synthesis we were able to do for this second part was limited and very context specific due to the limited nature and quantity of evidence available. The full results of this analysis will be reported in the full technical report, but we are unable to report on this here due to the lack of generalisable findings from this analysis and the limited space available here.

\(^4\) Specifically, we included: 1) Studies where participants are randomly assigned to treatment and comparison group (experimental study designs); 2) Studies where assignment to treatment and comparison group is based on other known allocation rules, including a threshold on a continuous variable (regression discontinuity designs) or exogenous geographical variation in the treatment allocation (natural experiments); 3) Studies with non-random assignment to treatment and comparison group, provided they include pre- and post-test measures of the outcome variables of interest to ensure equity between groups on the baseline measure, as well as use appropriate methods to control for selection bias and confounding, such as statistical matching (for example, propensity score matching, or covariate matching), regression adjustment (for example, difference-in-differences, and single difference regression analysis, instrumental variables, and ‘Heckman’ selection models).
3. Findings

3.1 Search and characteristics of the evidence base

We conducted a comprehensive search of the academic and grey literature and identified 66,430 citations for screening at title stage. Figure 2 provides the result of the literature search and screening at each stage. One thousand six hundred and seventy three papers were retained for screening at full text by two or more reviewers. Finally, 205 papers reporting on 151 studies and 142 unique interventions met our inclusion criteria and were included in the review.

3.2 Characteristics of included studies

As can be seen from figure 3 the included studies cover all major regions with L&MICs. We identified studies of programmes across 44 different countries, with 44 studies from Sub-Saharan Africa, 31 studies from East Asia & the Pacific, 46 from Latin America & the Caribbean; 34 from South Asia, one from Middle Eastern & North Africa and one from Europe & CIS.

Table 1 provides a list of all included interventions and the corresponding number of studies identified for inclusion. It highlights a divergence in the extent to which different interventions have been evaluated. It highlights few studies of teacher training, teacher incentives, teacher hiring and means based scholarships for instance. Other areas with few included studies are programs providing information either to children or parents about the benefits and returns to education, as well as grade retention, tracking, extra time and remedial education. Few studies evaluate the effect of providing materials (on their own), although this is an intervention component included in other programmes.

The most frequently reported outcomes were language test scores (n=159) and maths test scores (n=108). Reporting of cognitive outcomes was scarce (n=14) and was provided mainly for school-feeding and school based health studies. The most frequently reported access outcomes were enrolment (n=36), attendance (n=30) and drop-out (n=21).

Table 2 provides an overview of the study design used in included studies. As can be seen, the majority of studies used an experimental design. Of these, 53 per cent were cluster-randomised controlled trials and 9 per cent were randomised controlled trials, where random

---

5 The number of included papers is much larger than the number of included studies. This is because it is common for studies to be reported in more than one paper, typically one or more working papers and a journal version. Our approach was to make the most recent version the main paper, and then include any other version(s) of the paper which contained additional information. Typically the working papers and other ‘un-published’ reports would include more detail than journal versions, including effects on additional outcomes and sub-groups, and would therefore be included.

6 This number of unique interventions is lower as there were cases where we included several studies that use the same dataset but had different author teams undertaking different analysis or reporting additional outcomes.

7 We also include studies of cash transfers in our review, but due to the large size of the literature and the existence of relatively recent high quality SRs on the topic we’ve prioritised completing the synthesis for the other intervention areas first. The final report will include the analysis of these studies.

8 These numbers correspond to the number of included studies by outcome for those studies included in the meta-analysis only. Several other studies reported on these outcomes but were not suitable for synthesis.
assignment was done at the individual level. Almost a fifth of the studies used a controlled before and after design, with a smaller number of studies classified as cluster quasi-RCT, natural experiment and regression discontinuity designs.

3.3 Summary of findings on effects of interventions

The pooled estimates for all included primary outcomes are provided in figure 4 and this section summarises the results of the meta-analysis ordered by intervention interventions

School Feeding Interventions

We identified 11 studies that evaluated the effect of a school feeding programme implemented in Latin America and the Caribbean, Sub-Saharan Africa, East Asia and Pacific and South Asia.

The overall average effect calculated using random effects meta-analysis range from 0.01, 95% CI [-0.03, 0.01] for completion to 0.11, 95% CI [0.00, 0.21] for cognitive abilities. The meta-analyses of maths scores and language scores suggest effects that are smaller in magnitude for final learning outcomes (Maths: 0.04, 95% CI, [-0.2, 0.10]; Language: 0.04, 95% CI, [-0.01, 0.10]).

However, the homogeneity tests suggest a large amount of between study variability for nearly all outcomes, and individual effect sizes range from -0.9, 95% CI [-0.24, 0.07] in Peru to 0.37, 95% CI [0.27, 0.53] for grade 2 children in Senegal.

School-Based Health

We included 14 studies across seven countries in Latin America and the Caribbean, East Asia, South Asia and Sub-Saharan Africa that evaluated the effect of a school based health programme or trial in a low- or middle-income country.

The included studies evaluate a range of different school-based health interventions including de-worming (n= 5), malaria prevention and control (n= 4), micronutrient supplementation (n= 5), the provision of glasses (n= 1) and the provision of incentives for anaemia reduction (n= 1). Two studies provided a combination of de-worming and micronutrient supplementation.

The overall average effect range from 0.00, (95% CI [-0.23, 0.23]) for language test scores, to 0.08 (95% CI [-0.08, 0.25]) for maths test scores. The effect on school participation, as measured by attendance, was small 0.04 (95% CI [-0.08, 0.16]). Our results indicate that overall SBH interventions are beneficial, however the average effects are relatively small in magnitude and for some outcomes the effect is equal to zero. None of the meta-analyses indicated an overall negative impact of SBH on any outcome.

There is however considerable between study variability for all outcomes. To take one example, effect sizes for language test scores ranged from -0.26 (95% CI [-0.26, 0.56]) in Kenya to 0.56 (95% CI [0.39, 0.72]) in Sri Lanka (Fernando et al., 20069), even though both of these programs were malaria control interventions.

---

9 Fernando et al.’s (2006) study of a Malaria Control trial in Kenya was an outlier in terms of the large positive impacts that the programme had on language and maths tests scores (language: 0.56, 95% CI [0.39, 0.72]; maths: 0.62, 95% CI [0.45, 0.78]).
Merit-Based Scholarships

Merit-based scholarships aim to improve learning outcomes by rewarding high performing students with scholarships to continue their study (McEwan, 2013). We included five studies from East Asia and Pacific, South Asia, Sub-Saharan Africa, and Latin America and the Caribbean.

The overall average effect range from 0.01, 95% CI [-0.14, 0.15], for attendance to 0.11, 95% CI [0.04, 0.18] for maths. The overall average effect of merit based scholarships on language scores is 0.04 SMD (95% CI [-0.07, 0.15]), 0.10, 95% CI [0.03, 0.18]) for composite test scores. There is however considerable between study variability for all outcomes.

User Fee Elimination

We included seven studies that evaluated the effect of programmes which reduced school user fees, either by abolishing all or parts of school fees. The studies covered programmes in Latin America and the Caribbean, Sub-Saharan Africa and East Asia and Pacific.

The overall average effect ranges from a reduction in attendance of -0.05, 95% CI [-0.09, -0.02] to an improvement in drop-out rates of -0.10, 95% CI [-0.23, 0.02]10. The effect on enrolment is positive but small in magnitude (SMD = 0.04, 95% CI [0.01, 0.08]), whereas the meta-analysis suggest an overall reduction in completion rates (SMD=-0.03, 95% CI [-0.13, 0.06]). None of the studies report on effects on learning outcomes.

There is however considerable between study variability for all outcomes and the meta-analysis is based on few studies, so the results should be interpreted with caution.

Scholarships

Programmes that provide scholarships and allowances to households aim to cover all or some of the costs associated with education, including school fees, uniforms and books. Such scholarships are different from merit-based scholarships in that they are paid regardless of educational performance or ability. We included four studies that evaluated the effect of five programmes that provided scholarships to students in South Asia and East Asia and Pacific.

The overall average effect for attendance suggest a small positive effect (SMD=0.02, 95% CI [0.02, 0.03]), whereas the meta-analysis of drop-out outcomes suggest a small increase in drop-out rates among students receiving the scholarship as compared to those that did not (SMD= 0.03, 95% CI [-0.08, 0.13]). The overall effect on maths is not different from zero (SMD=0.00, 95% CI [-0.11, 0.12]).

However, there is a large amount of between study variability and not more than three studies reporting on any one outcome, so the results should be interpreted with caution.

Computer-Assisted Learning

Computer-assisted learning have been widely implemented as a means of tailoring learning to students’ individual needs (Kremer, 2013). We identified fourteen studies of the effect of twelve unique CAL programmes implemented in a range of settings in East Asia and the Pacific, South Asia, and Latin America and the Caribbean.

The overall average effect range from -0.04, 95% CI [-0.11, 0.04] for enrolment to 0.05, 95% CI [0.00, 0.11] for composite test scores. There is also a large amount of between study

10 Note that a negative effect on drop-out favours the intervention.
variability for all outcomes. Individual effect sizes range from a -0.36, 95% CI [-0.44, -0.27] for OLPC on language arts in Peru (Cristia et al., 2011) to a SMD of 0.20 [0.12, 0.24] for maths test scores among children in a CAL program in China. The large amount of heterogeneity is not surprising given the variability in the intervention design, context and populations in our included studies.

Based on the studies included in our review it is not clear that the overall effect of CAL on children’s learning is beneficial. While all average effects are small, for several outcomes they are negative and the average effect hides studies with negative effects that are quite large in magnitude.

Providing Materials for Teaching and Learning

Providing materials such as books, chalkboards or other classroom equipment are often claimed to assist teachers to facilitate learning and improve educational quality. We identified four studies of programmes providing materials across three different countries in South Asia and Africa.

The overall average effect ranges from -0.02, 95% CI [-0.06, 0.02] for maths, to 0.01, 95% CI [-0.01, 0.02] for composite test scores. It suggests that children in schools receiving materials or grants for the purpose of buying materials do not do any better than children in schools not receiving such materials. While there is a large amount of heterogeneity, there is a similar pattern of small or negative effects across learning outcomes.

This finding is based on few studies and should be interpreted with caution. Many of the other interventions assessed at the school level do suggest beneficial effects of programmes that include the provision of materials together with other components. One interpretation of this result may therefore be that in addition to issues with poor implementation, materials is a necessary but not sufficient condition for children's learning.

Remedial Education

Remedial education refers to a range of interventions aimed at improving learning outcomes for students who are lagging behind their peers, or normative standards of achievement. Three studies assess the effect of remedial education on learning outcomes in three programmes in Chile, Mexico and India. We were able to examine effects on maths, language arts and composite test scores using meta-analysis.

The overall average effect for maths is 0.07, 95% CI [0.05, 0.08]. The average effect on language arts and composite test scores is also positive, if slightly smaller in magnitude (SMD=0.04, 95% CI [0.02, 0.06] and SMD=0.06, 95% CI [0.04, 0.08]). There is some between study variability, but all effect estimates have a positive sign, with only one study having confidence intervals that cross the line of no effect. The individual study effects range from 0.02 [-0.02, 0.05] (Banerjee et al., 2007) for the full sample in Mumbai (year 1) to 0.10 for several of the Grade 4 sub-samples in Vedodara. The limited number of studies precludes any strong conclusions about the effect of remedial education on learning outcomes.

New Schools and Infrastructure

New schools and infrastructure interventions aim to improve enrolment and attendance rates and, thereby, also improve learning outcomes over time.

We included six new schools and infrastructure studies implemented in South Asia, Sub-Saharan Africa, Latin America and the Caribbean, and Europe and Central Asia. Due to the diversity of interventions and outcomes reported in this category we were only able to conduct meta-analysis of effects of hygiene infrastructure on enrolment and attendance.
The results of the meta-analysis suggest hygiene infrastructure inventions such as construction of latrines can have a beneficial effect on school participation (SMD enrolment=0.11 SMD (95% CI [0.01, 0.20]; SMD attendance=0.14, 95% CI [0.05, 0.24]). However, because of a high amount of between study heterogeneity and limited evidence base the results should be interpreted with caution.

**Pedagogy interventions**

Pedagogy interventions include initiatives that aim to adapt or improve educational content, or the instructional approach by which this content is taught, typically in combination with new materials for children. We identified twenty studies of pedagogy programmes across eleven different countries in South Asia, South-East Asia, Latin America and Sub-Saharan Africa. The overall average effects on participation outcomes range from 0.02, 95% CI [0.00, 0.04] for attendance to 0.13, 95% CI [0.02, 0.28] for completion. Access is not a key objective of most pedagogy interventions so it is unsurprising that these interventions do not seem to have affected attendance rates in participant schools.

Overall average effects relating to learning range from 0.01, 95% CI [-0.04, 0.07] for cognitive test scores, to 0.06, 95% CI [0.03, 0.08] for composite test scores, 0.11, 95% CI [0.03, 0.18] for maths test scores and 0.24, 95% CI [0.12, 0.36] for language arts test scores. There is a relatively large amount of between study variability. Effect sizes for maths test scores range from -0.09, 95% CI [-0.15, -0.03] to 0.40 95% CI [0.34, 0.47], while effect sizes for language test scores range from -0.02, 95% CI [-0.10, 0.07] to 0.87 95% CI [0.48, 1.25].

**Extra Time in School**

Some educational programmes are introduced to increase the amount time students spend in school by extending the school day. We identified two studies, from Chile and Ethiopia, evaluating such interventions. The overall pooled effect for language arts is relatively large in magnitude (SMD=0.21, 95% CI[0.12, 0.30]), with a slightly smaller effect for maths (SMD=0.09, 95% CI[-0.04, 0.22]), although this is based on only two studies.

**School participation by ability**

We identified two studies assessing different interventions using student ability to allocate students to classes and groups. Chen et al (2010) evaluate the effect of grade retention in China, while Duflo et al. (2011) study the effect of tracking in Kenya. Due to the heterogeneity in these interventions we were not able to conduct a meta-analysis.

**Teacher incentives**

Teacher incentive interventions seek to improve the working conditions in schools so that teachers are motivated to come to work and improve their performance. Such interventions take many forms, such as providing direct payments to teachers based on their attendance or on the achievement of their students, and teacher surveillance and monitoring.

We identified eight studies of teacher incentive programmes across five different countries in Latin America and the Caribbean, Sub-Saharan Africa and South Asia. The overall average effect ranges from 0.00, 95% CI [-0.13, 0.12] for language arts to 0.03, 95%CI [0.01, 0.05] for maths. There are few studies reporting on these outcomes but among those that do, the findings of homogeneity tests indicate that there is no or small between-study variability and where relevant, results of sensitivity analysis indicate that these results are not sensitive to the removal of any of the included effect sizes.

**Teacher Hiring Interventions**
Interventions to improve teacher hiring aim to ensure that posts in new schools can be filled, existing schools can expand, or teacher-pupil ratios can be increased. These initiatives tend to focus on the conditions of contracts of teachers (e.g. being hired on a permanent basis or on contract).

We identified seven studies of teacher hiring programmes across five different countries in South Asia, Latin America and Sub-Saharan Africa. The overall average effect of receiving an additional contract teacher range from 0.04, 95% CI [0.01, 0.08] for completion, to 0.06, 95% CI (-0.01, 0.12) for composite test scores, 0.07, 95% CI (0.03, 0.10) for language arts test scores and 0.10, 95% CI (0.00, 0.20) for maths test scores. These results should be interpreted cautiously as they are based on just a few studies.

**School-Based Management Interventions**

School-based management (SBM) interventions include initiatives which de-centralise authority to the school level. SBM programmes may involve handing decision-making on budget, staffing and curriculum development over to teachers, parents, students or other community members (Barrera-Osorio, 2009). We identified eleven studies of SBM across six different countries in Latin America, East Asia and Sub-Saharan Africa. The overall average effects range from 0.00, 95% CI [-0.09, 0.09] for completion to 0.06, 95% CI [-0.02, 0.15] for language arts. While this result indicates that, overall, SBM are beneficial as compared to the control, the average effects are relatively small in magnitude for all outcomes.

There is also a large amount of between study variability for all outcomes, limiting the generalisability of the overall average effect. Individual effect sizes range from a -0.12 [-0.24, -0.0] for teacher attendance in COGES programme in Niger (Beasley et al. 2014) to a SMD of 0.32 [0.09, 0.55] for completion among children in the Plano de Desenvolvimento da Escola (PDE) programme in Brazil (Carnoy et al. 2008).

**Community-Based Monitoring**

Community based monitoring (CBM) interventions seek to provide information and improve the representation of communities in which service providers, governments, or other public bodies operate (Westhorp et al., 2013). We identified eleven studies of CBM programmes across seven countries in South Asia, Sub-Saharan Africa and Latin America and the Caribbean. The overall average effect ranges from 0.04 for student attendance (95% CI [-0.1, 0.18]) to 0.17 for student enrolment (95% CI [0.08, 0.25]). However, there was significant heterogeneity in results for all participation outcomes, and all the meta-analyses apart from that on enrolment are based on few studies.

The combined average effect of CBM initiatives on learning outcomes ranges from 0.07 (95% CI [0.01, 0.13]) for composite test scores, to 0.09 for maths test scores (95% CI [-0.02, 0.2]) and 0.12 (95% CI [0.01, 0.22]) for language test scores. There is however significant heterogeneity within the outcomes across programs and the average estimate should be interpreted with caution.

**Private-Public Partnerships**

Public private partnerships and private provision of schooling (PPP) may seek to increase parents’ and students’ choice, provide a supply of schooling when there is none, or improve the quality of education provided (Barrera-Osorio et al., 2009). We identified eleven studies of PPP programmes across six different countries in South Asia and Latin America. The overall average effect ranges from 0.03, 95% CI [-0.03, 0.10] for enrolment to 0.05, 95% CI
[-0.05, 0.15] for composite test scores. While this result indicates that overall PPP are beneficial, the average effects are relatively small in magnitude for all outcomes.

There is also a large amount of between study variability for all outcomes, suggesting the effects of PPP may differ between contexts. Individual effect sizes range from a -0.04 [-0.08, 0.00] for language arts in the Andhra Pradesh School Choice Project in India (Muralidharan and Sundararaman, 2013) to a SMD of 0.20 [0.16, 0.24] for composite test scores among children attending private schools through a voucher program in Colombia.

4. Discussion and conclusions

Our review of the evidence on the effects of twenty different education programs suggest that while most interventions have an overall positive effect on beneficiaries as compared to children not receiving these interventions, for many intervention areas the effects are very small in magnitude, and for some outcomes our analysis also find zero or small negative effects.

Nevertheless, we are able to identify some programmes that appear particularly promising. Assessing the relative effects of the twenty different interventions included in our review we find that pedagogy interventions have the largest average effect on learning outcome. The meta-analysis for language arts outcomes includes twenty studies, many of them large scale RCTs, and the meta-analysis show a SMD of 0.24 (95% CI:0.12, 0.36]). The overall effect on maths test scores is slightly smaller in magnitude (SMD=0.11, 95% CI [0.03, 0.18]). But as most of the programmes evaluated focus on literacy in particular this is perhaps not surprising.

Interventions increasing the time children spend in school by extending the school day also show an effect which is larger in magnitude compared to the other intervention types. The overall pooled effect is 0.21, 95% CI[0.12, 0.30] for language arts outcomes and 0.09, 95% CI[-0.04, 0.22]) for maths, although this is based on only two studies. Nevertheless, these studies evaluate large scale programs in Chile and Ethiopia.

Looking at the results of the meta-analyses on participation outcomes we find that the effect of CBM on school enrollment is larger in magnitude than what we find for other interventions (SMD=0.17, 95% CI [0.08, 0.25], 12 studies). It is the third largest in magnitude across all comparisons in the review and the largest relative effect on enrollment. However, the size of the effect is not consistent across outcomes, with smaller effects on attendance, drop-out and completion. Effects on learning outcomes range from 0.07, 95% CI [0.01, 0.13] for composite test scores, to 0.12, 95% CI [0.01, 0.13] for language arts.

For other participation outcomes we observe the largest relative effect on school attendance for new schools and infrastructure programmes (latrines, SMD=0.14, 95% CI [0.05, 0.24], 2 studies). For reduction in drop-out rates we find the largest effect for elimination of school user fees (SMD=-0.10, 95% CI [-0.23, 0.02], 4 studies). Finally, for completion we find the largest relative effect for pedagogy interventions (SMD=0.13, 95% CI [-0.02, 0.28], 2 studies). All of these findings are base on few studies however.

As noted in the section above, we observe a large amount of varaibility for most pooled effect estimates. This means that the differences in effects observed across programmes are not due to chance, but rather due to other factors, such as programme design, implementation, context and populations in the included studies. But unfortunately, due to the limited number of included studies for any single meta-analysis we were not able to conduct any moderator analysis to assess which factors may be driving the results.
Our attempts at conducting qualitative synthesis to identify factors potentially explaining the results were limited by the quantity and quality of available evidence. Few impact evaluations report on such factors and the findings from associated qualitative studies and process evaluations are typically too context specific to allow for any generalisable findings at the intervention or review level.

Finally, while we were able to identify a large number of high quality impact evaluations, the usefulness of many studies for evidence synthesis is at times compromised by the quality of reporting. Details on context, intervention design and implementation is often missing, so it can be difficult to assess what was delivered, by whom and at what cost. Reporting of details on sample characteristics, as well as exact sample size, standard errors and/or confidence intervals is also lacking in many studies, resulting in a large burden on review authors in trying to get this data, and failing that, the exclusion of studies from the review.

Acknowledgements: The systematic review is funded by the International Initiative for Impact Evaluation (3ie). The team want to acknowledge contributions from Dafni Skaldiou and Shari Krishnaratne on earlier versions of the protocol. Hugh Waddington and Dr. Howard White provided helpful comments on the protocol. The team also gratefully acknowledge the feedback from our Advisory group members Dr. Robert Slavin and Dr. Frances Hunt. Finally, want to acknowledge valuable consultancy services on effect size data extraction and meta-analysis provided by Dr Ariel Aloe, Chris Thompson and Shannon Shisler.
Figure 2: PRISMA diagram

78,238 records identified through academic database searching

1034 records identified through other sources [grey literature search, citation tracking, targeted search etc.]

66,430 records screened at title (after duplicates removed)

60,785 records excluded

7116 records screened at abstract

5571 records excluded

1673 full-text articles assessed for eligibility

Exclude on date: 1
Exclude on country: 10
Exclude on population: 63
Exclude on intervention: 177
Exclude on study design: 424
Exclude on outcome: 282
Exclude on efficacy: 76
Exclude as duplicate of included article: 72

205 impact evaluations included in review preliminary

Corresponding to 151 studies and 142 unique interventions preliminary

Figure 3: Geographical coverage of included studies
Table 1: Number of studies by intervention type

<table>
<thead>
<tr>
<th>Intervention Level</th>
<th>Intervention Category</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>System</td>
<td>Community-Based Monitoring</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Public-Private Partnerships</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>School-Based Management</td>
<td>11</td>
</tr>
<tr>
<td>Teacher</td>
<td>Teacher Hiring</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Teacher Incentives</td>
<td>8</td>
</tr>
<tr>
<td>School</td>
<td>Computer Assisted Learning</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Pedagogy</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Extra Time</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>New Schools and Infrastructure</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Providing Materials</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Remedial Education</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Grade Retention</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Tracking</td>
<td>1</td>
</tr>
<tr>
<td>Household</td>
<td>Providing Information to Parents</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Reducing/Eliminating User Fees</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Scholarships</td>
<td>4</td>
</tr>
<tr>
<td>Child</td>
<td>Merit Based Scholarships</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Providing Information to Children</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>School-Based Health</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>School Feeding</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 2: Overview of included study designs

<table>
<thead>
<tr>
<th>Study design</th>
<th>Number</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomised Controlled Trial (RCT)</td>
<td>12</td>
<td>9%</td>
</tr>
<tr>
<td>Cluster RCT</td>
<td>73</td>
<td>53%</td>
</tr>
<tr>
<td>Cluster-quasi RCT</td>
<td>11</td>
<td>8%</td>
</tr>
<tr>
<td>Natural experiment</td>
<td>8</td>
<td>6%</td>
</tr>
<tr>
<td>Regression Discontinuity Design (RDD)</td>
<td>8</td>
<td>6%</td>
</tr>
<tr>
<td>Controlled Before and After (CBA)</td>
<td>25</td>
<td>18%</td>
</tr>
<tr>
<td>Interventions</td>
<td>Enrolment</td>
<td>Attendance</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Child level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School feeding</td>
<td>0.03 [-0.01, 0.06]</td>
<td>0.09 [-0.02, 0.20]</td>
</tr>
<tr>
<td>School-based health</td>
<td>No studies</td>
<td>0.04 [-0.09, 0.16]</td>
</tr>
<tr>
<td>Merit-based scholarships</td>
<td>No studies</td>
<td>0.01 [-0.14, 0.15]</td>
</tr>
<tr>
<td>Providing information to children</td>
<td>No meta-analysis</td>
<td>No meta-analysis</td>
</tr>
<tr>
<td><strong>Household level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User fee elimination</td>
<td>0.03 [-0.02, 0.07]</td>
<td>-0.05 [-0.09, -0.02]</td>
</tr>
<tr>
<td>Scholarships</td>
<td>No meta-analysis</td>
<td>0.02 [0.02, 0.03]</td>
</tr>
<tr>
<td><strong>School level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer-assisted learning</td>
<td>-0.04 [-0.11, 0.04]</td>
<td>0.04 [0.00, 0.07]</td>
</tr>
<tr>
<td>Providing materials</td>
<td>No meta-analysis</td>
<td>No meta-analysis</td>
</tr>
<tr>
<td>Remedial education</td>
<td>No studies</td>
<td>No studies</td>
</tr>
<tr>
<td>New schools and infrastructure</td>
<td>0.11 [0.01, 0.20]</td>
<td>0.14 [0.05, 0.24]</td>
</tr>
<tr>
<td>Pedagogy</td>
<td>No studies</td>
<td>0.02 [0.00, 0.04]</td>
</tr>
<tr>
<td>Extra time in school</td>
<td>No studies</td>
<td>No studies</td>
</tr>
<tr>
<td><strong>Teacher level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher incentives</td>
<td>No meta-analysis</td>
<td>0.01 [-0.04, 0.06]</td>
</tr>
<tr>
<td>Teacher hiring</td>
<td>No studies</td>
<td>No meta-analysis</td>
</tr>
<tr>
<td><strong>System level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School-Based management</td>
<td>0.01 [-0.04, 0.07]</td>
<td>No meta-analysis</td>
</tr>
<tr>
<td>Community-based monitoring</td>
<td>0.17 [0.08, 0.25]</td>
<td>0.04 [-0.10, 0.18]</td>
</tr>
<tr>
<td>Private-public partnerships</td>
<td>0.03 [-0.03, 0.10]</td>
<td>No studies</td>
</tr>
</tbody>
</table>
5. References:

**Included impact evaluations**


Bellei C (2009) Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. Economics of Education Review. 28: 629-640.


Bjorkman M (2006) Does Money Matter for Student Performance? Evidence from a Grant Program in Uganda...


Blimpo Moussa P; (2010) Team Incentives for Education in Developing Countries A Randomized Field Experiment in Benin. : Unpublished working paper.


Cabezas V, Cuesta J I; Gallego F A; (2011) *Effects of Short-Term Tutoring on Cognitive and Non-cognitive skills: Evidence from a Randomized Evaluation in Chile*. Cambridge, MA: JPAL.


Carlos David, Laura Quispe (n.d.) *Evaluación del Impacto del Programa OLPC sobre el Rendimiento Académico en el Área de Matemática*. Perú: Universidad Católica Santa María.


Chen Zhejiang Xinxin; Shi Yaojiang, Yi Hongmei, Zhang Linxiu, Mo Di, Chu James, Loyalka Prashant, Rozelle Scott (2013) *The Impact of a Senior High School Tuition Relief Program on Poor Junior High School Students in Rural China*. Xi’an: Northwest University.


Fernando Deepika, de Silva , Damani , Carter Richard, Mendis Kamini N; Wickremasinghe Rajitha (2006) A randomized, double-blind, placebo-controlled, clinical trial of the impact of


He Fang, Linden Leigh L; MacLeod Margaret (2007) Helping Teach What Teachers Don’t Know: An Assessment of the Pratham English Language Program. New York: Columbia University.


http://www.hkazianga.org/Ppapers/KaziangadeWalqueAldermanEducation.pdf


Lai F, Zhang L, Qu Q, Hu X, Shi Y, Boswell M, Rozelle S ( ) Teaching the Language of Wider Communication, Minority Students and Overall Educational Performance: Evidence from a Randomized Experiment in Qinghai Province, China.


Lara Bernardo, Mizala Alejandra, Repetto Andrea (2009) Revisiting the School Choice Debate in Chile.


Linden L L; (2008) *Complement or substitute? the effect of technology on student achievement in India*. : Columbia University.


Loyalka Prashant, Liu Chengfang, Song Yingquan, Yi Hongmei, Huang Xiaoting, Wei Jianguo, Zhang Linxiu, Shi Yaojiang, Chu James, Rozelle Scott (2013) Can information and


Maiga Eugenie (2010) *Essays on the Economics of Health and Education in Developing Countries*.


Mo Di, Linxiu Zhang, Qinghe Qu, Matthew Boswell, Scott Rozelle (2013) Integrating Computer Assisted Learning into a Regular Curriculum: Evidence from a Randomized Experiment in Rural Schools in Shaanxi.


Orkin K (2013) *Lengthening the school day and children’s achievement in Ethiopia*: Unpublished manuscript


Rau T B; Contreras D G; (2011) *Tournaments Incentives for Teachers: The Case of Chile*. Unpublished manuscript.


Saavedra Facusse, T B (2013) *Efecto del financiamiento compartido sobre el rendimiento escolar*.


**Other references:**


