

How context mediates the effectiveness of teacher support interventions: Evidence from a multi-arm randomised experiment in South Africa¹

Stephen Taylor²

Paper submitted for the Economics Society of South Africa (ESSA) conference of 2017

August 2017

Abstract

A major constraint to economic development is the low quality of school education provided in many countries. The body of research on what works to improve learning outcomes in schools is growing rapidly, with a large proportion of evaluated interventions being some form of teacher training or support. However, for maximum usefulness to policy makers looking for sustainable teacher support interventions on a large scale, it is crucial to understand where and why different types of support work. Contextual factors may well play a decisive role in whether a programme has any positive impact. This forms the basis of a common external validity critique of randomised experiments.

This paper reports on the Early Grade Reading Study, which is a randomised experiment conducted amongst low socio-economic status schools in South Africa, with three separate interventions all aimed at improving the teaching and learning of Home Language literacy in the early grades. Two interventions involved pedagogical support to teachers using structured lesson plans, but differed in the modality of training the teachers. A third intervention aimed to improve parent involvement in home literacy activities. All three interventions had large positive impacts on literacy outcomes in urban township settings, but all three had negligible effects in deep rural settings.

The paper explores the reasons for this dramatic difference in the impact of interventions using a variety of methods. This analysis investigates socio-economic and other differences between urban and rural areas, changes in intermediate outcomes including classroom practice, and monitoring data on the fidelity of programme implementation. Some of the emerging factors prohibiting impact in rural settings include long distances which prevent attendance at programme activities, and lost teaching time (perhaps due to weaker accountability of schools to government and parents).

¹ This paper draws on the Early Grade Reading Study (EGRS), which is being led by the Department of Basic Education in partnership with the Human Sciences Research Council (HSRC), the University of the Witwatersrand and various donors (3ie, UNICEF, ZENEX Foundation, Anglo American, and Department of Planning, Monitoring and Evaluation). The full Research Team comprises Stephen Taylor, Brahm Fleisch (WITS), Jacobus Cilliers (Georgetown University), Cas Prinsloo (HSRC) and Vijay Reddy (HSRC).

² Department of Basic Education, government of South Africa; Research Associate of the Department of Economics, University of Stellenbosch

Background to the EGRS

One of the biggest developmental challenges facing South Africa is the high number of children who do not learn to read for meaning in the early years of school. This is the foundational skill upon which all others build and as such this has become a leading priority for the Department of Basic Education (DBE). In order to address this challenge, the DBE initiated the Early Grade Reading Study (EGRS) in collaboration with academics at the University of the Witwatersrand, the Human Sciences Research Council (HSRC) and Georgetown University (USA). This is a large-scale educational impact evaluation – the biggest in South Africa - and aims to build evidence about what works to improve the teaching and learning of early grade reading in African languages in the country.

Learning to read is foundational to all subsequent learning; yet the majority of South African children are being left behind in this regard. South Africa's participation in international assessments of reading and literacy such as PIRLS (2006, 2011) and SACMEQ (2007, 2013) has revealed that large proportions of children in grades 4, 5 and 6 have not yet learned to read with comprehension. These children, who have not *learned to read*, can therefore not *read to learn* in subsequent grades and in all their subjects.

Based on all the available evidence it is clear that weak reading foundations are the major root cause of weak academic achievement later in school and subsequently the dropping out of school that occurs amongst 16 to 18 year-olds in South Africa. It is widely accepted that performance in the secondary school leaving examination, known as “matric”, is strongly predictive of post-schooling success and wages earned in the labour market. Furthermore, numerous economists have shown that wage inequality is the main driver of overall economic inequality in South Africa. Therefore, it is not an exaggeration to say that the most effective way to bring about meaningful socio-economic transformation in South Africa is to improve the learning and teaching of reading in schools serving historically disadvantaged communities.

Although there are various initiatives underway to support early grade reading – and there have been many others in the past, there is little or no evidence of what is working or why. A Randomised Control Trial (RCT) design provides an accurate estimation of the *causal* impact of interventions, and thus has the potential to inform responsible policy decisions. By using a random lottery to allocate schools to intervention and control groups (no intervention) it is possible to construct a credible “counterfactual” scenario i.e. what would have happened to those who did receive the intervention if they had not received the intervention.

Project Design

The core of the project is a comparison of the cost-effectiveness of three promising intervention models to improve reading outcomes in learners' home language (Setswana). The project commenced in 2015 by working in 230 schools in quintiles 1-3 in the North West province. Each intervention has been implemented in a separate group of 50 schools with an additional 80 control schools where ordinary schooling is continuing. The project uses a formal impact evaluation methodology known as a Randomised Control Trial (RCT) complemented with a 60-classroom observation study and eight detailed case-studies. The study design enables the researchers to

estimate the impact of each intervention model on measures of reading, as well as understand where, how and why different elements of the intervention models are working.

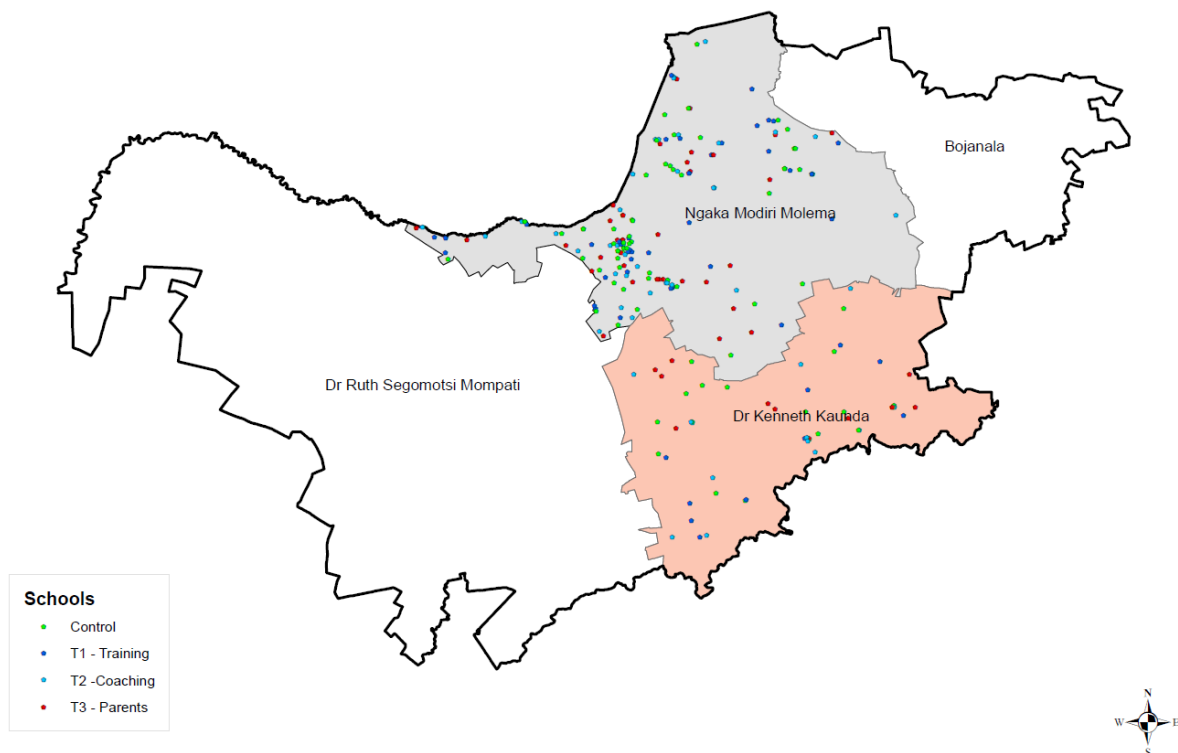
The evaluation assessed three intervention models:

1. **A structured learning programme & centralised training:** The first intervention provides teachers with lesson plans aligned to the National Curriculum Statement Grades R-12 (NCS) including the Curriculum and Assessment Policy Statements (CAPS), as well as additional quality reading materials and training at centralised workshops twice a year.
2. **A structured learning programme & specialist on-site coaching:** The second intervention (implemented in a different group of 50 schools) provides teachers with the same set of lesson plans and reading materials but provides ongoing support to teachers through on-site coaching and small cluster training sessions.
3. **Parental intervention:** The third intervention (implemented in a further 50 schools) holds weekly meetings with parents to discuss the importance of learning to read in the early grades and to empower them with the knowledge and tools to become more involved in their child's literacy development.

The three interventions were implemented in the grade 1 class of 2015 and at the grade 2 level in 2016, thus following the same cohort of learners. This year (2017) the two structured pedagogic interventions have continued at the grade 3 level, therefore ensuring that this cohort of learners are exposed to the interventions for the entire Foundation Phase. An organisation called "Class Act" has served as the service provider for the implementation of the interventions.

Over the same period of time, **three waves of data on the same learners** have been collected by the HSRC. The institutional independence between the implementation of interventions and the evaluation side of the project is a distinctive feature of the EGRS. A baseline data collection ("Wave 1") was collected at the start of 2015 when learners had just begun grade 1. A midline data collection ("Wave 2") was collected at the end of 2015 where learners were close to completing grade 1. A third wave of data was collected at the end of 2016, when most learners were in grade 2. Data collected towards the end of 2016 when the learners had received two years of the interventions forms the basis for the evaluation findings presented in this report. At each wave of data collection, 20 randomly selected learners per school were individually assessed on Setswana reading and literacy. The Early Grade Reading Assessment (EGRA) was used for certain items whilst other items were added, with slight changes occurring between each wave to accommodate the changing skill levels of learners as they progress through school. At each data collection point, we also administered a school principal questionnaire, teacher questionnaires and parent/guardian questionnaires, which were sent home with learners and returned to the school. This background information is useful for analysis of differential programme effects for various subgroups of learners, teachers and schools, and for examining whether teacher or parent attitudes and practices may have shifted in response to the interventions.

Figure 1: Map of North West province showing the random allocation of schools to each of the different interventions



Literature review: Early Grade Reading interventions internationally and in South Africa

Improving the quality of education in developing countries has been a conundrum that has received significant attention over the past two decades. Consequently, there has been a surge in the implementation of various interventions aimed at affecting change, and commensurately so, the evaluation of these interventions. A broad range of interventions have been tried and evaluated, including the provision of information about the quality of schooling to parents, providing in-service training to teachers, providing additional resources to schools, providing new technology to classrooms, teachers or learners, implementing nutritional and health interventions, changing the incentive and accountability structures in which teachers operate, and providing cash transfers or merit-based scholarships.

The first generation of school based experiments mainly focused on increasing school resources, but found that these inputs did not produce the expected achievement gains. The provision of improved school infrastructure has been found to affect school attendance, but has not made a significant impact on learner performance (Adukia, 2017; Freeman, et al., 2011). Learning and teaching support material (LTSM) are often assumed to be essential in supporting learning, however, the mere distribution of these resources has not proven to be successful in improving learning outcomes (Glewwe, Kremer, & Moulin, 2009; Glewwe, Kremer, Moulin, & Zitzewitz, 2004; Das, Dercon, Habyarimana, Krishnan, Muralidhara, & Sundararaman, 2011; Sabarwal, Evans, & Marshak, 2014). The effectiveness of resources such as LTSM appears to depend on how well these are used by teachers, and even by factors such as whether learners are able to read those

materials. These studies, however, have mainly focused on the distribution of resources, without additional support or training in the use of these resources.

A different strand of interventions focused on improving learner health and nutrition with the purpose of enabling learners to attend school more regularly and to learn more productively. Nutrition and health interventions are thought to affect learning outcomes through improving learner attendance and learner concentration. These interventions typically include programmes that administer either deworming medication or nutritional supplements to learners, running a school feeding scheme that provides a nutritious meal to learners, the provision of reading glasses, providing immunisations and malaria prevention programmes. School feeding programmes have shown to have a small effect on learning outcomes, but a larger effect on learner attendance especially in areas where food security is low (Diagne, Lô, Sokhna, & Diallo, 2014; Ismail, Jarvis, & Borja-Vega, 2012; Altman, 2013; McEwan P. , 2013). The evidence of deworming programmes is largely inconclusive, with some evaluations of deworming finding some educational benefits for learners (Ebenezer, et al., 2013; Miguel & Kremer, 2004), but evaluations in other contexts not finding any positive impact (Watkins, Cruz, & Pollitt, 1996; Simeon, Grantham-McGregor, Callender, & Wong, 1995). Similarly, programmes that aim to provide micronutrient supplements to learners had beneficial effects on learners in some contexts (Luo, et al., 2012; Wong, Shi, Luo, Zhang, & Rozelle, 2014), but not in others (Jukes, Zuilkowski, Parawan, & Lee, 2014).

The lack of accountability of schools and teachers for professional conduct and providing quality education is often cited as a reason for weak learner performance in developing countries. For instance, high levels of teacher absenteeism and the fact that teacher pay is typically unrelated to any performance measures are often cited as evidence of this. This has led to experiments with interventions that aim to improve education outcomes through parents or communities holding schools accountable. Similar to the health and nutrition interventions, the results of increasing accountability measures vary greatly and seem to be largely dependent on the context. Positive results have been found for interventions that provide parents and community members with more information on the oversight role that they can play, or provide parents with school score cards (Andrabi, Das, & Khwaja, 2015; Pandey & Goyal, 2011). However, in different contexts the provision of information to parents did not lead to any increased community involvement or teacher effort (Banerjee, Banerji, Duflo, Glennerster, & Khemani, 2010; Nguyen & Lassibille, 2008).

A significant body of research has shown that teachers and their teaching are critical to learner performance (e.g. Hanushek, 2010). Various interventions therefore aim to improve the quality of teaching, either through teacher training, teacher incentives, changes to how teachers are hired or providing diagnostic feedback to teachers. Evidence seems to suggest that teachers do not necessarily change their instructional practices in response to financial incentives, but that change may occur if participation in a training programme has explicit implications for promotions or salary increases (Glewwe, Illias, & Kremer, 2010; Murhaldiharan & Sundararaman, 2011; Popova, Evans, & Arancibia, 2016). In a systematic review of 26 teacher in-service training programmes, Popova et al. (2016) found that the largest impacts on student learning come from programmes that focus primarily on classroom management. However, their results also suggest that more generic training programmes without any focus on a specific subject have a lower impact on student learning (Popova, Evans, & Arancibia, 2016). This review also showed that face-to-face training sessions at universities or training centres had larger impacts on learner performance than sessions held at centralised venues such as government buildings or hotels, although the former programmes may typically be more time-intensive. Programme impact was also found to be improved if follow-up

visits which involved reviewing the material that was taught in the initial training was included in the programme. Researchers or government officials were also found to be less effective trainers than education practitioners (Popova, Evans, & Arancibia, 2016).

Programmes which seem to have consistently shown positive and large impacts on learner performance are those where resources are provided as part of a more comprehensive intervention package which includes training and support in the use of these resources (Piper, Zuilkowski, & Mugenda, 2014; Nonoyama-Tarumi & Bredenberg, 2009). In a systematic review conducted by Snilstveit, et al. (2016), programmes that combine teacher training and resources are classified as **structured pedagogy programmes**. In their review, Snilstveit, et al. identified 21 different structured pedagogy programmes and found that these programmes are effective in improving learner performance in both mathematics and reading, but that larger impacts were observed in the language programmes. Studies that found particularly high gains include the Primary Maths and Reading Rural Expansion programme in Kenya and the School Readiness Programme in Cambodia (Piper, Zuilkowski, & Mugenda, 2014; Nonoyama-Tarumi & Bredenberg, 2009).

In South Africa, three studies measuring the impact of reading programmes on learner performance are worth noting. The first programme, in 2000, was the Learning for Living project which was implemented by the Read Educational Trust (READ). The programme provided resources to classrooms and mentoring to teachers. The evaluation took on a quasi-experimental design, where READ staff members nominated schools that had achieved high levels of programme implementation as the treatment schools to be evaluated. Government officials further selected schools that closely mirrored the demographics of the intervention schools as the control schools. This design was of course unlikely to sufficiently control for selection effects and this remains a limitation. The evaluation was conducted four years into the implementation and assessed grade 1 and grade 2 learners in their home language and found strong positive effects for the READ Home Language Initiative (Sailors, Hoffman, Pearson, Beretvas, & Matthee, 2010).

The second study piloted a programme called the Systematic Method for Reading Success (SMRS), which used both lesson plans and the Early Grade Reading Assessment (EGRA) tool, was conducted by the Molteno Institute of Language and Literacy in collaboration with the Department of Education, and was evaluated by RTI International. The programme was conducted in ten treatment and five control schools in each of the Limpopo, Mpumalanga and North West provinces in South Africa during 2009. The evaluation found that over a five month period, the programme increased the average letters read per minute from 1.75 per minute to 16.09 per minute (Piper, 2009). The small sample size of this study, however, makes it difficult to gain a good understanding regarding the scalability of this programme.

The final study on early grade reading in South Africa is the Gauteng Primary Literacy and Maths Strategy (GPLMS) that was implemented in the Gauteng province in 2010. The core components of the strategy included daily lesson plans, high-quality learning and teaching materials and ongoing instructional coaching. The study was evaluated using a Regression Discontinuity Design, but suffers from various limitations in the identification of a control group and with the outcomes data. Notwithstanding these limitations, both the implementation and the evaluation of GPLMS suggested that a high quality structured learning programme supported by instructional coaching could be effective at a relatively large scale (Fleisch, Schöer, Roberts, & Thornton, 2016).

The international and South African literature was influential in shaping the design of interventions chosen for evaluation in the EGRS. Both sets of literature pointed to the promise of pedagogic learning programmes, but left critical questions unanswered: To what extent did the modality and dosage of teacher support influence the effectiveness of the programme? What would be the impact of such programmes after three years of implementation in typical South African Foundation Phase classrooms? On parent involvement, would it be possible to shift learning outcomes at low cost through engaging parents in South Africa, as had been achieved in other contexts such as Pakistan (Andrabi, Das, & Khwaja, 2015)?

Description of EGRS Interventions

Intervention 1 (“Training”) - A structured learning programme & centralised training

The first intervention provides teachers with daily lesson plans aligned to the National Curriculum Statement Grades R-12 (NCS) including the Curriculum and Assessment Policy Statements (CAPS), as well as additional quality reading materials and training at centralised workshops twice a year.

The four learning areas in the curriculum for grades 1 to 3 are Home Language literacy, First Additional Language (which is usually English), Numeracy, and Life Skills. The lesson plans are thus intended to strengthen the enactment of the curriculum and should not be seen as an alternative to current policy. They provide detailed specifications for each lesson including information on methodology and content to be taught for each instructional day. The lesson plans incorporate the use of learning support materials including the government-provided “DBE workbooks” as well as certain additional materials (graded reading booklets, flash cards and posters), which are provided through the EGRS. The graded reading booklets provide a key resource for the teacher to use in group-guided reading and individual work so as to facilitate reading practice at an appropriate pace and sequence of progression. EGRS provided the Setswana “Vula Bula” graded reading book series developed by the Molteno Institute for Language and Literacy. These books were developed in the relevant African languages as opposed to being translated, and progress in accordance with the natural phonic progression of each language.

Intervention 1 trains the teachers on how to use the lesson plans and accompanying materials through central training sessions, each lasting 2 days, and occurring twice a year. These sessions were conducted for grade 1 teachers in February and July of 2015 and for grade 2 teachers in January and July of 2016. The table below details the teaching and learning materials that teachers were provided with during the project.

Table 1: Intervention 1 materials

| ITEM | DESCRIPTION |
|-----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Vula Bula Reading Books | Commercially produced Grade One and Grade Two Setswana graded reading books. These were used in group guided reading lessons |
| Book register | An exercise book set up as an accession register for the Vula Bula reading books |
| Teacher file | A management file to keep teaching and learning materials |
| Setswana HL scripted lesson plans | This document contains the individual lesson plans that teachers followed in 2015 and in 2016 |
| Flashcard words | Printed sets of the words teachers needed to teach sight words in reading lessons |
| Reading words | A learner resource that listed the sight words taken from the Vula Bula books. These word lists were taken home so that the learners could practice reading |
| Assessment records | CAPS and SA-SAMS compliant assessment record tables. Teachers used this resource to record formal assessments per learner |
| Assessment rubrics | Criteria for teachers to use to award objective assessment ratings for learner tasks |
| Curriculum tracker | A tool for teachers to manage curriculum coverage |
| Weekly routine | A tool for teachers to manage curriculum pacing |
| Core methodologies | Detailed pedagogical support that helped teachers learn how to use tried and tested methodologies for different language components |
| Handwriting poster | A poster that demonstrated the form and directionality of lower and upper case letters |
| Theme posters | Posters that detailed interesting scenes that were used for vocabulary development |
| Facilitators' Guides | Detailed handbooks for trainers to follow when they trained teachers. |

The 2-day training events occurred four times over the course of 2015 and 2016 and were well attended, as can be seen in Table 2. The first training event in February 2015 covered the lesson plans for Term 2 only (and not Term 1) since the learning programme only began in Term 2 due to the other preparatory activities taking place in Term 1. This would not have meant any disruption to the learning programme in intervention schools since the lesson plans are aligned to the official curriculum, only specifying a greater level of detail and with particular activities, instructional methods and resources integrated into the lessons. Only on one occasion was a school not represented at all at a training session. The attendance rate for teachers was also high (between 85% and 100%) and was sustained throughout the two years. Teachers who did not attend the residential training sessions were provided with catch-up training. Attendance rates were a little lower for school leaders (principals or HODs) but this is not a major concern since they were not the primary recipients of training.

Table 2: Attendance rates at training events

| | GRADE ONE (2015) | | GRADE TWO (2016) | |
|---------------------------|------------------------------|-----------------------------------|----------------------------------|-----------------------------------|
| | ACTUAL TERM 2 FEB 2015 | ACTUAL TERM 3 & 4 JULY 2015 | ACTUAL TERM 1 & 2 JAN 2016 | ACTUAL TERM 3 & 4 JULY 2016 |
| % Schools attended | 100 | 98 | 100 | 100 |
| % Teachers attended | 100 | 85 | 98 | 93 |
| % School leaders attended | 74 | 78 | 90 | 80 |

Intervention 2 (“Coaching”) - A structured learning programme & on-site coaching

The second intervention (implemented in a different group of 50 schools) provides teachers with the same set of lesson plans and reading materials but provides ongoing support to teachers through on-site coaching and small cluster training sessions. Therefore, if the lesson plans are implemented with the same level of fidelity across Interventions 1 and 2, classroom practice and hence learning outcomes should be identical across the two groups. However, the modality of supporting teachers differs. Instead of bi-annual central training sessions, ongoing support to teachers consisting of regular (monthly) on-site coaching from “reading coaches” is provided. In addition to these on-site visits, there are occasional meetings with the coach and a small cluster of teachers from nearby Intervention 2 schools. The evaluation of Interventions 1 and 2 should thus shed light on a) whether the structured pedagogy programme can improve the enactment of the curriculum and thus improve reading acquisition, and b) whether the mode of teacher support is important in determining effective enactment.

Table 3 shows how the coaches were allocated to schools. The schools were divided geographically across the three coaches. Many schools had more than one teacher to support.

Table 3: Allocation of reading coaches

| COACH | DISTRICT | GRADE ONE (2015) | | GRADE TWO (2016) | |
|---------|---------------------|-------------------|------------------|-------------------|------------------|
| | | NUMBER OF SCHOOLS | GRADE 1 TEACHERS | NUMBER OF SCHOOLS | GRADE 2 TEACHERS |
| Coach 1 | Ngaka Modiri Molema | 17 | 27 | 17 | 26 |
| Coach 2 | Ngaka Modiri Molema | 18 | 34 | 18 | 32 |
| Coach 3 | Dr Kenneth Kaunda | 14 | 34 | 14 | 31 |
| TOTAL | | 49 | 95 | 49 | 89 |

Table 4 summarises the attendance of teachers at the various training engagements as well as the dosage of on-site coaching visits. High attendance levels were noted throughout the project demonstrating ongoing commitment. Teachers were supported throughout the project in their classrooms between 2 and 3 times per term. Fewer coaching visits were possible per teacher in the last term of 2016 due to a combination of social unrest in one district and learner assessments and other outside disruptions in both districts. In addition to classroom-based support, teachers received additional support during needs-driven afternoon workshops amongst nearby clusters of schools, which were facilitated by coaches. Although these workshops did happen to some extent in Year 1, these support initiatives became more structured in Year 2 and were therefore reported on in Year 2. Due to the shortened length of Term 4 and disruptions to schooling in the area, no afternoon workshops were run by coaches.

Table 4: Summary of attendance and dosage of Intervention Two

| | GRADE ONE (2015) | | | GRADE TWO (2016) | | | |
|--------------------------------------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | TERM 2 FEB 2015 | TERM 3 JUL 2015 | TERM 4 SEP 2015 | TERM 1 JAN 2016 | TERM 2 APR 2016 | TERM 3 JUL 2016 | TERM 4 SEP 2016 |
| % Schools attended 1-day training | 100 | 92 | 100 | 100 | 100 | 96 | 100 |
| % Teachers attended 1-day training | 100 | 89 | 100 | 99 | 100 | 92 | 99 |
| Average number of on-site coaching visits | 3 | 2 | 2 | 2 | 3 | 3 | 1 |
| % Teachers attended cluster-based afternoon workshops | | | | 48 | 59 | 61 | 0 |

Intervention 3 – Parental involvement

The third intervention (implemented in a further 50 schools) holds weekly meetings with parents to discuss the importance of learning to read in the early grades and to empower them with the knowledge and tools to become more involved in their child's literacy development.

At each school in this group, a Community Reading Coach (CRC) was recruited. The CRC was identified through communication with the school principal who recommended a suitably qualified and available person in the community. The CRCs attended a 1-day training session facilitated by the service provider (Class Act) at the start of each school term (quarterly). The CRCs were trained to deliver weekly training sessions for parents at their respective schools. For their services, CRCs were paid a stipend of R400 per month (about \$35). Under this arrangement, CRCs were essentially volunteers receiving a small stipend, rather than employees receiving a salary.

A total of 30 sessions were scheduled for each year covering a total of 10 topics per year. Each topic had 3 sessions where the topic was the same but the activities of the session differed. Thus, a parent could attend roughly 1 in 3 sessions and still be exposed to all topics, while parents who attended more regularly could still enjoy fresh activities. The topics covered in these sessions included the importance of learning to read for later educational and labour market success, training on how to support their child's reading at home and the provision of low-cost materials and reading games to use at home. As with Interventions 1 and 2, grade 1 parents were invited in 2015 and grade 2 parents in 2016.

Table 5: Intervention 3 materials

| ITEM | DESCRIPTION |
|----------------------|-------------------------------------------------------------------------------------------------------------|
| Module One | 'Small things can make a difference' plus a set of family reading cards |
| Module Two | 'Playing with sounds to support reading' plus a set of family reading cards |
| Module Three | 'Reading pictures' plus a set of family reading cards |
| Module Four | 'Letter sounds' plus a set of family reading cards |
| Module Five | 'Incidental reading' plus a set of family reading cards |
| Module Six | 'Preparing to read a story – Part One' plus a set of family reading cards |
| Module Seven | 'Reading a story' plus a set of family reading cards |
| Module Eight | 'Preparing to read a story – Part Two' plus a set of family reading cards |
| Module Nine | 'Reading Remediation' |
| Module Ten | 'Reading stories 3 and 4' |
| Module Eleven | 'Reading story 5' |
| Module Twelve | 'Reading stories 6, 7 and 8' |
| Module Thirteen | 'Reading stories 9, 10 and 11' |
| Module Fourteen | 'Reading stories 12, 13 and 14' |
| Module Fifteen | Consolidation |
| Grade One reader | Platinum Series Le Re Tlhabetse Readers published by Maskew Miller Longman / Pearson: Book 1 |
| Grade Two reader | Platinum Series Le Re Tlhabetse Readers published by Maskew Miller Longman / Pearson: Book 2 |
| Facilitators' Guides | Detailed handbooks for trainers to follow when they trained CRCs. These were also used for parent training. |

Summary of Year 1 (Midline) Evaluation Results

A full Midline technical report was compiled and the results can be summarised as follows. We observed small to moderate positive impacts of both the Training and the Coaching interventions on Setswana reading outcomes at the end of grade 1. Under the assumption that the gains from baseline to midline in the control group reflect a year of learning, we estimate that the two pedagogical interventions yielded an impact of approximately 20% of a year of learning during three quarters of a year.

Overall, the impact of the parent involvement intervention was small, and a zero impact could not be ruled out. The most likely reason for low average impact was low attendance rates amongst parents at the weekly meetings.

The impacts of the Training and the Coaching interventions were clearer for boys than for girls. For boys, each of these interventions had an estimated effect of 0.19 standard deviations, and this was statistically significant at the 95% level of confidence. This result could be a positive finding for the sake of helping boys catch up to girls in literacy outcomes.

The positive effects of the Training and the Coaching interventions were clearer amongst schools in urban areas (33% of our sample), where the estimated effects were higher than 30% of a standard deviation. We also observed that children in the upper middle range of achievement stood to benefit most from the pedagogical interventions, possibly indicative of the level at which the lesson plans were set. However, no children experienced negative impacts.

We also observed some evidence of shifts in intermediate outcomes in the form of changed teacher and classroom practice. We found that grade 1 teachers in Interventions 1 and 2 were more likely to “stream” children into groups according to their reading proficiency, compared to the control group. Intervention 2 teachers appeared to conduct individualised reading assessments of learners more frequently than the control group. There was some evidence of increased reading resources in Intervention 1 and 2 classrooms, especially of Setswana posters. Encouragingly, based on an inspection of learner exercise books, there was consistent evidence of more exercises of all types (including drawing pictures), of written exercises, and of full sentence writing exercises in both Intervention 1 and 2 schools compared to the control group.

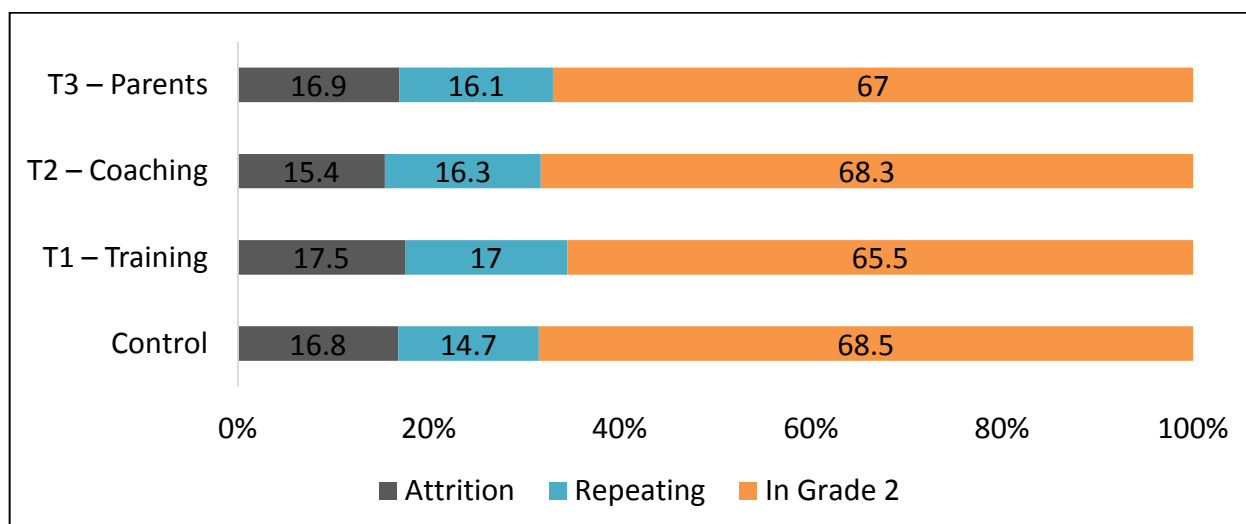
Attrition and Grade repetition between Waves 1, 2 and 3

At baseline we assessed 4 538 learners at the start of their grade 1 year in 2015. At the midline assessment (end of grade 1) we successfully re-assessed 4 143 learners, meaning that about 9% of the original sample were either absent from school on the day of the survey, had moved to another school, had returned to Grade R, or had stopped attending school. However, at the end of grade 2, just over half of those absent for the midline were successfully re-assessed, implying that they must have been absent from school on the day of the midline survey. Therefore, we can say that about 4% of the original sample appeared to have left the school before the end of grade 1.

At the time of the Wave 3 data collection (end of grade 2, 2016), we successfully assessed 3 781 learners (83.3% of the original sample). This means that, over and above the 4% of learners who left the school before the end of grade 1 in 2015, a further 13% were either absent on the day of the Wave 3 survey or had left the school since grade 1. Unfortunately, we are not able to distinguish between these two reasons.

Amongst those learners who were successfully identified at the Wave 3 survey, 591 were found to be repeating grade 1, and these learners were tested again. Figure 2 shows the proportions of the original sample that attrited, that were repeating grade 1, and that were found to be in grade 2. Importantly, there were no significant differences in attrition or grade repetition across the three intervention groups. This means that the tested samples of learners will not be selectively stronger or weaker in any intervention group, something that could have introduced bias into the impact analysis.

Figure 2: Attrition and grade repetition in the sample



Year 2 evaluation findings: What works?

The learner assessment administered at the end of grade 2 (October/November 2016) was designed primarily to measure home language (Setswana) literacy outcomes, as this was the primary goal of each of our interventions. However, we included two grade-appropriate mathematics items and a few English reading items since these reflect the other two main learning areas within the Foundation Phase, namely mathematics and English as a First Additional Language. Several teachers in the project had spoken about how the home language learning programme we were running was time-consuming and difficult to fit into the school timetable. Therefore, we suspected that it was possible that the

EGRS home language literacy programme could have a negative effect on either mathematics or English through crowding out instructional time in those learning areas. Alternatively, we hypothesised that an effective home language programme could have positive spillover effects on mathematics or English through the acquisition of transferable skills or through the teacher's practice in one subject area improving due to training applied to another area. Therefore, we decided to include the mathematics and English items in the learner assessment.

Components of learner assessment

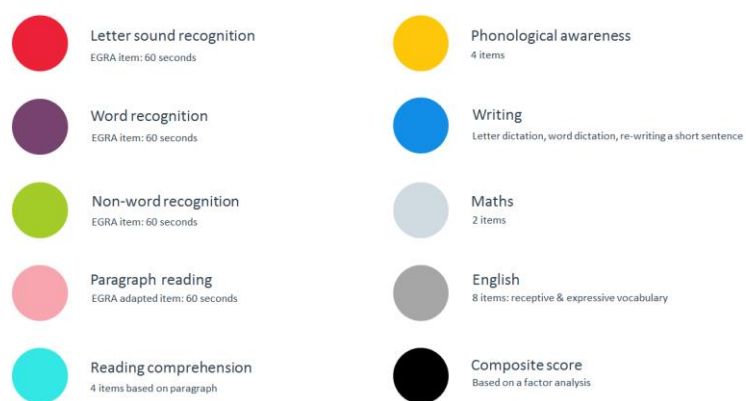
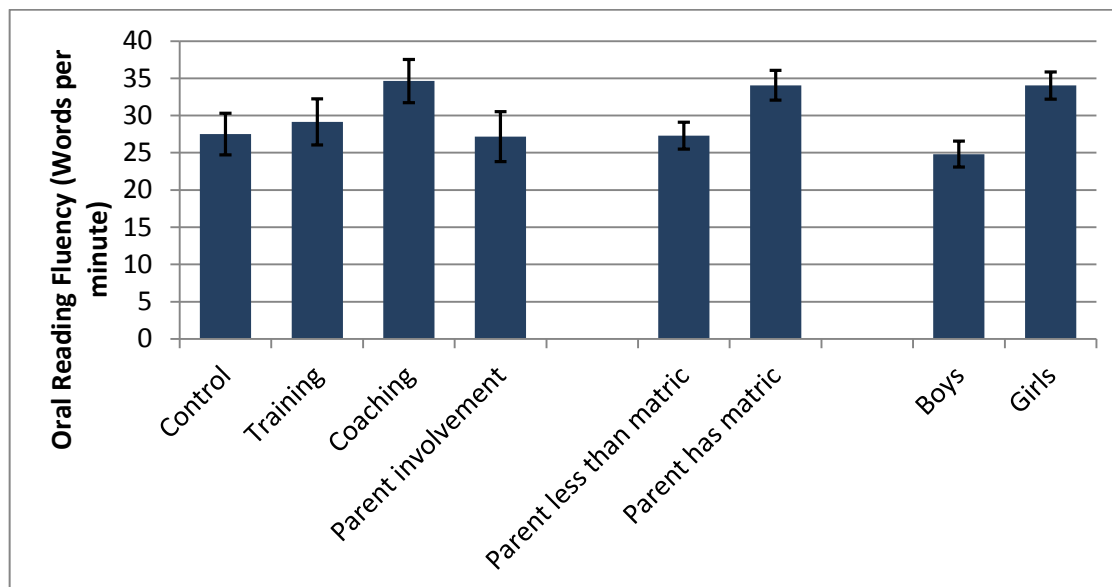


Figure 3 reports how the different groups of children performed in Oral Reading Fluency (reading aloud) by the end of grade 2. It is evident that learners in the Coaching group were furthest ahead of the control group by the end of grade 2. Those in the Coaching group could read approximately 7 more words per minute than those in the control group – amounting to 26% more words per minute. Children whose parents had at least a matric performed substantially better than those whose

parents were less educated, indicating the importance of the educational environment at home. Girls performed substantially better than boys, reading nearly 10 words per minute more than boys on average. This is a worldwide finding - that girls have better early grade reading results than boys, but the gaps are particularly large in South Africa and persist throughout schooling – ultimately girls are more likely to complete secondary school than boys.

Figure 3: Average Oral Reading Fluency by subgroups of interest at the end of Grade 2



Note: Learners repeating Grade 1 are not included in this figure; 95% confidence intervals are indicated.

The next three figures present further descriptive evidence of the differences in achievement between the Coaching group and the control group. Figure 4 shows the percentage of children achieving above particular thresholds of words correct per minute, separately for the two groups of children. The scores at the 25th percentile, median (50th percentile) and 75th percentile of the distribution for the full sample of 230 schools are also indicated on the graph. In both groups there were roughly 85% of children who could read at least one word correctly. There were also similar percentages of children who managed to read all 50 words correctly within a minute (about 6%). However, throughout the range between zero and 100% there were consistently more Intervention 2 children able to surpass particular thresholds. Between the thresholds of about 10 words per minute and 25 words per minute there were consistently about 10% more children in the Intervention 2 group able to read at least that number of words than in the control group. The pattern in this graph points to the possibility that the impact of the coaching intervention was largest for children in the mid-range of the performance distribution.

Figure 4: Word recognition for Intervention 2 and Control

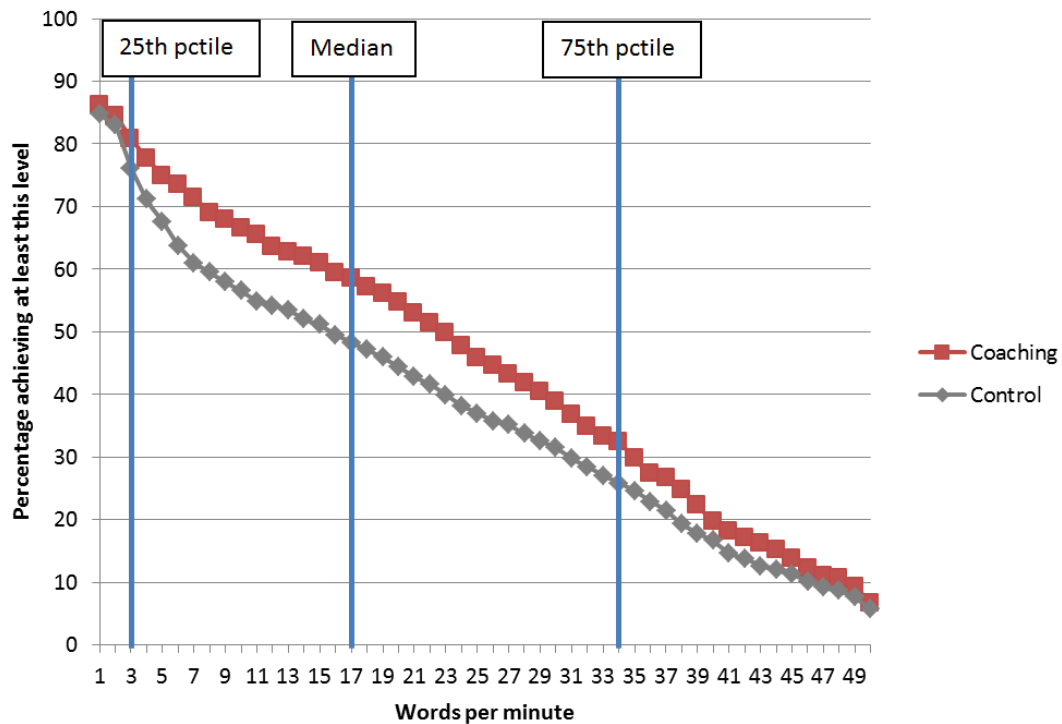


Figure 5 below shows the same type of graph as the one above, now applied to the Paragraph reading test, which provides a measure of oral reading fluency. As before, similar proportions of children could complete reading the entire paragraph within a minute (about 10%). A floor effect on this item (36% of children scoring zero) means that it is difficult to say anything about the impact on the bottom end of the distribution; however, we can say that there were about 72% of children in the Coaching group who could read at least one word correctly, whereas only 61% of children in the Control group could do so. Similarly, about 10% more children surpassed the median level of achievement (23 words per minute) in the coaching group compared to the control group. Only for the top 30% of the distribution did the magnitude of this impact drop off.

Figure 5: Oral Reading Fluency for Intervention 2 and Control

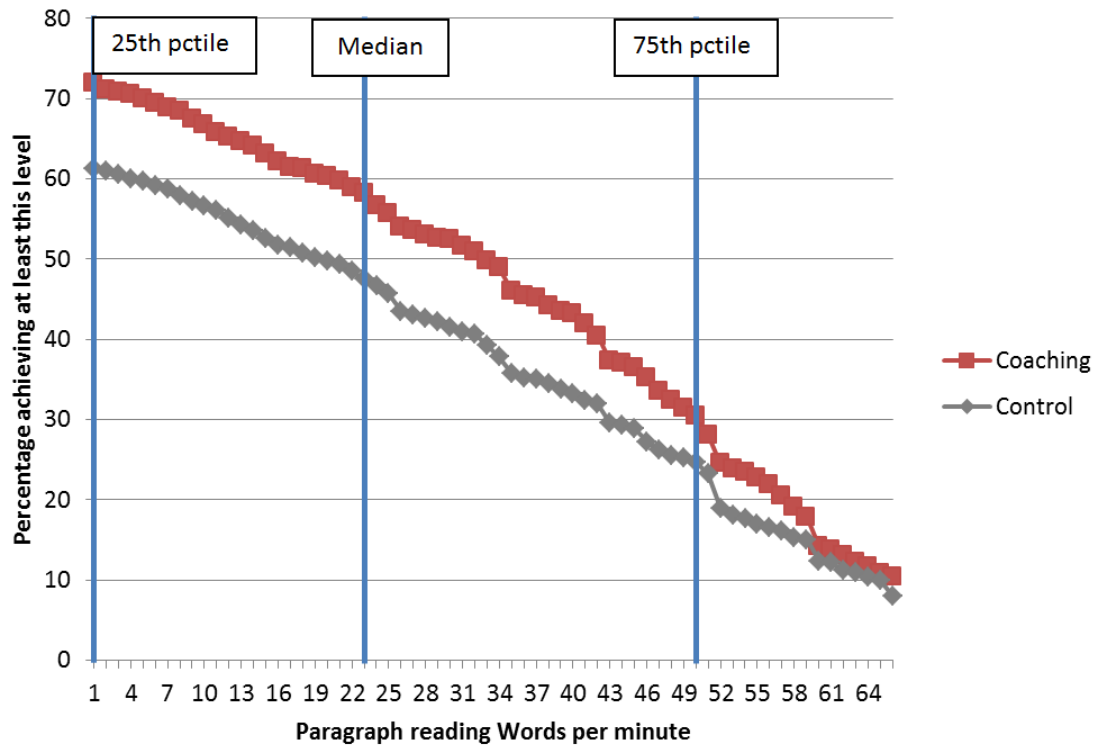
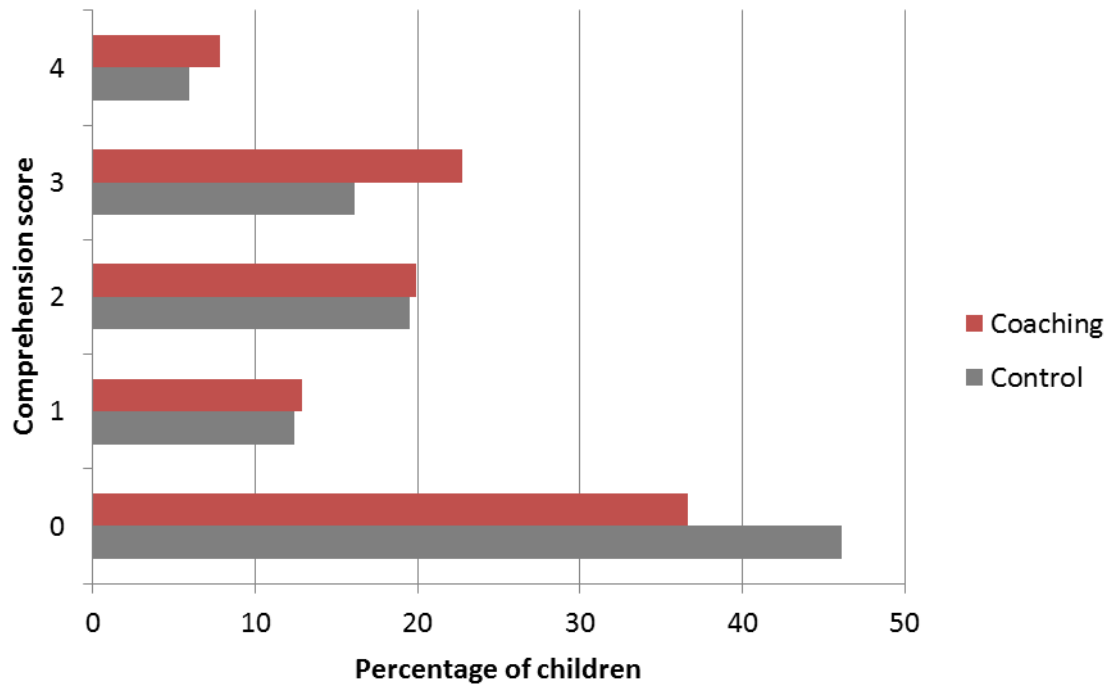


Figure 6 shows the percentages of children achieving each possible score out of 4 on the comprehension test, which was administered after the paragraph reading. Whereas 46% of children in the control group scored zero, only 37% of those in the coaching group scored zero. The entire distribution shifted upwards for the intervention group, with about 10% more children scoring either 3 or 4 out of 4 than in the control group.

Figure 6: Comprehension scores for Intervention 2 and Control



The descriptive analysis of the shifts in performance for the Coaching group is useful to get a sense of the magnitudes of the effects that will be described in the forthcoming regression analysis. Table 6 presents the results from our preferred model specification, which controls for baseline scores, district (schools are spread randomly across two districts), school mean score in the Annual National Assessments of 2014 (the most recent standardized school assessment), learner gender, parent education (according to the parent/guardian questionnaire), and two community-level controls obtained from the national census of 2011, namely a community wealth index derived from several questions about household possessions and the proportion of 13 to 18 year-olds in the community that are attending and educational institution. The motivation for including these controls is to account for any incidental differences that may exist between the treatment groups as well as to improve the precision of the estimates by increasing the explanatory power of the model. Only the coefficient on the Coaching treatment group is statistically significant at conventional levels, while the estimated effects of Training and Parent Involvement are small and not statistically significant.

Table 6: Year 1 regression models with full controls (including repeaters)

| | Intervention 1 (Training) | Intervention 2 (Coaching) | Intervention 3 (Parents) |
|----------------|------------------------------|------------------------------|-----------------------------|
| Intervention 1 | 0.112 (0.0814) | | |
| Intervention 2 | | 0.252*** (0.0792) | |
| Intervention 3 | | | 0.103 (0.0768) |
| Constant | -1.601** (0.624) | -1.596** (0.674) | -1.066* (0.591) |
| Observations | 2,121 | 2,140 | 2,140 |
| R-squared | 0.170 | 0.178 | 0.183 |

Notes: Standard set of controls included
Cluster robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

When excluding repeaters the estimated impacts of Interventions 1 and 2 increases, as one might expect given that only grade 2 teachers received additional support in 2016. We can now be 90% sure that Intervention 1 (training) had a non-zero positive impact. However, the effect size is perhaps somewhat on the small side given that this was after two years of intervention, and we cannot be 95% sure that the impact was not zero. The effect of Intervention 2 (coaching) on those children who received two years of the programme is estimated to be 0.33 standard deviations, and we can be 99% sure that there was a non-zero impact. The 90% confidence interval ranges from about 0.2 standard deviations to about 0.5 standard deviations. Relative to the RCT literature on educational interventions, this is certainly a substantial effect size.

Table 7: Year 1 regression models with full controls (excluding repeaters)

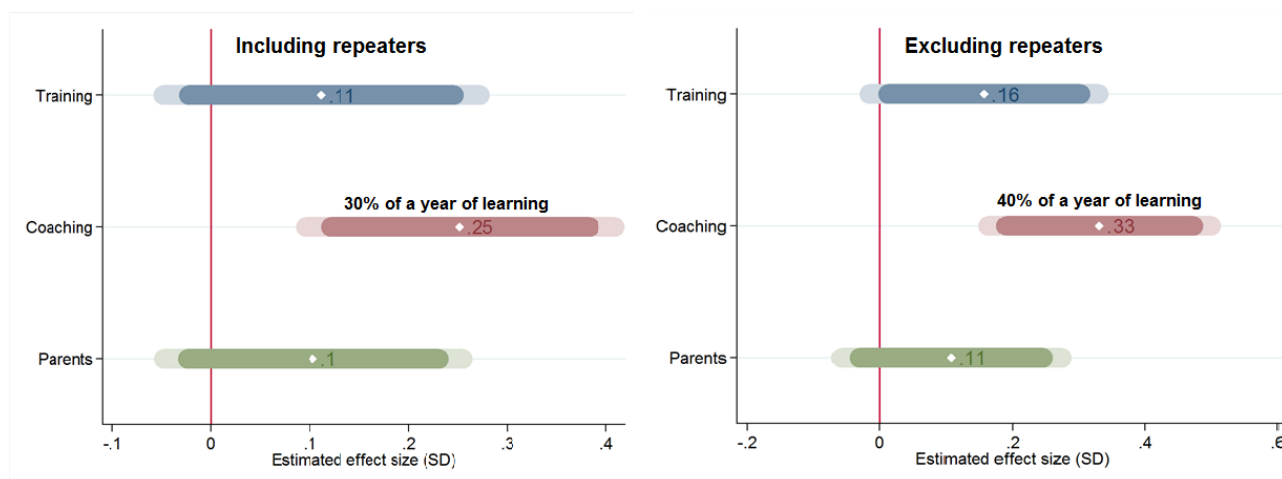
| | Intervention 1 (Training) | Intervention 2 (Coaching) | Intervention 3 (Parents) |
|----------------|------------------------------|------------------------------|-----------------------------|
| Intervention 1 | 0.158* (0.0875) | | |
| Intervention 2 | | 0.332*** (0.0853) | |
| Intervention 3 | | | 0.108 (0.0840) |
| Constant | -1.276* (0.695) | -1.450** (0.695) | -0.775 (0.680) |
| Observations | 1,758 | 1,772 | 1,781 |
| R-squared | 0.163 | 0.180 | 0.178 |

Notes: Standard set of controls included
Cluster robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 7 graphically presents the estimated effect of each intervention relative to the control group, based on a multivariate regression analysis where the outcome is a composite home language reading score. Effects are expressed in standard deviations of test scores gained.

Of the three intervention models, the Coaching intervention shows a substantial positive impact after two years of intervention (end of grade 2). **Learners who received two years of this Coaching intervention were approximately 40% of a year of learning ahead of the students in the schools that received no intervention ('business-as-usual' schools).** This is a truly significant improvement by international standards. The other two interventions (centralised training and the parent involvement intervention) appeared to have a small positive impact, less than half the size of the Coaching intervention. If we include learners repeating grade 1 (about 16% of the sample), who only received the first year of the interventions, the average effects are somewhat smaller as indicated in the left-hand pane of Figure 7.

Figure 7: Estimated intervention effects relative to the control group based on regression models



Note: The graphs show the point estimates of each intervention's effect in standard deviations of test scores, 90% confidence intervals (dark shaded bars), and 95% confidence intervals (light shaded bars).

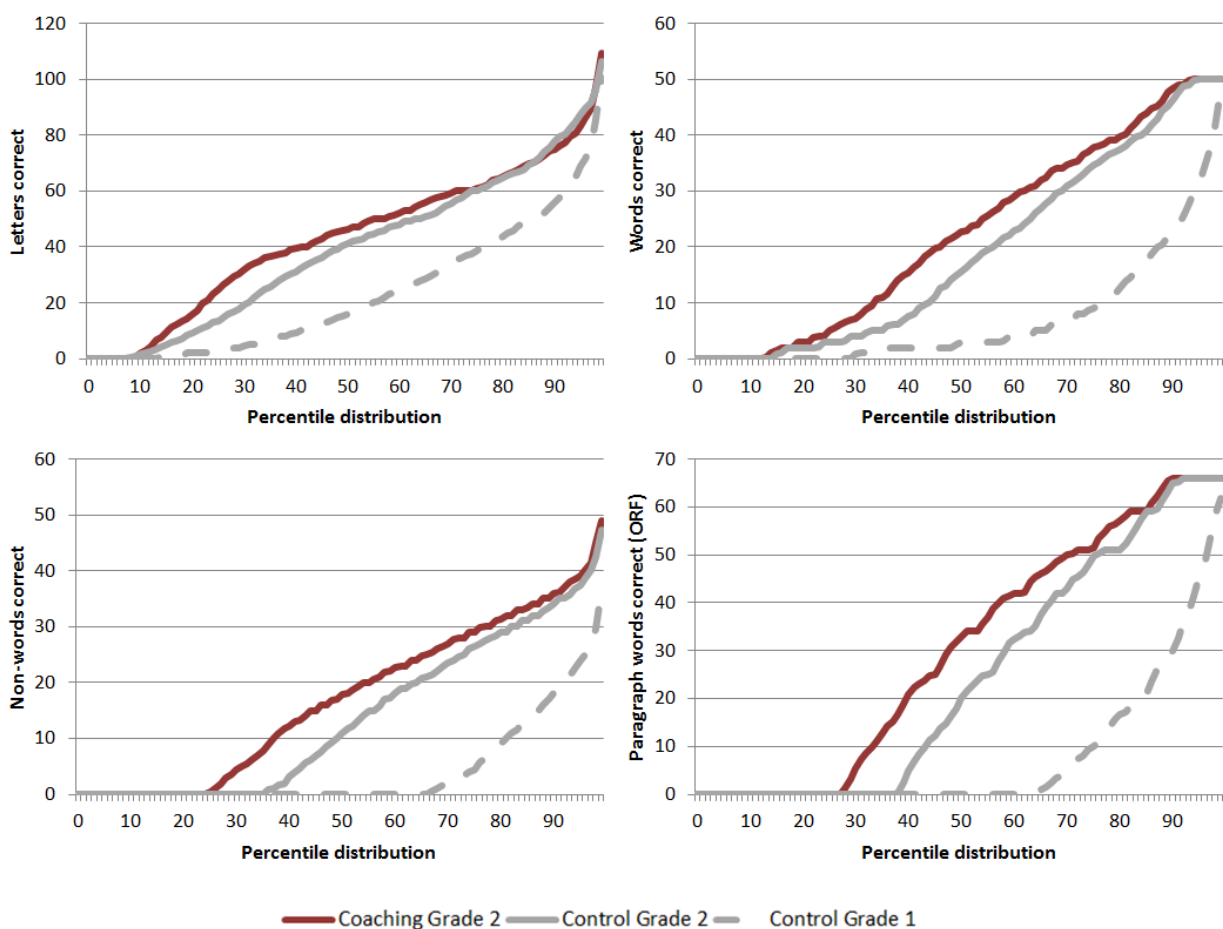
Intervention effects relative to a year of learning

It is common practice to estimate intervention effects in terms of standard deviations of test scores using multivariate regression models, which take other characteristics predictive of test scores into account. When effect sizes are reported in terms of standard deviations this has some advantages, but the educational significance of such measurement units are not necessarily transparent, especially to audiences not versed in statistics.

One way to provide an intuitive measure of an intervention's impact is to compare the intervention effect size to the amount of learning that would normally occur in a year of schooling. We can think of the change in test scores in the control group over 1 year as reflecting a year of learning. In order to do this we use only those items which were common across both Waves 2 and 3. The extent to which learners in the control group improved on these items between the end of grade 1 (Wave 2) and the end of grade 2 (Wave 3) provides an estimate of 1 year of learning. The idea is then to see how much extra learning the intervention groups achieved and express this additional amount as a percentage of 1 year of learning in the control group.

However, this exercise is not as simple as it may sound, as a consideration of Figure 8 demonstrates. The figure shows percentile plots (the score at each point in the percentile distribution of performance) for those items common to Waves 2 and 3. The gap between the line for the control group at the end of grade 1 and the line for the control group at the end of grade 2 represents a “year of learning”. The gap between the control group at the end of grade 2 and the Coaching group at the end of grade 2 represents the impact of the intervention, i.e. the additional learning over and above the counterfactual. For example, one year of learning (at the 50th percentile) in terms of letter recognition is estimated to be 25 letters (the gap between 15 letters at the end of grade 1 and 40 letters at the end of grade 2). At the 50th percentile of grade 2 learners in the coaching group, the score was 46 letters correct, which was 6 letters more than the control group. Therefore, it could be estimated that the impact on letter recognition was 24% of a year of learning $((6/25)*100)$ – at least at the 50th percentile.

Figure 8: Percentile plots for sub-tests common across Waves 2 and 3



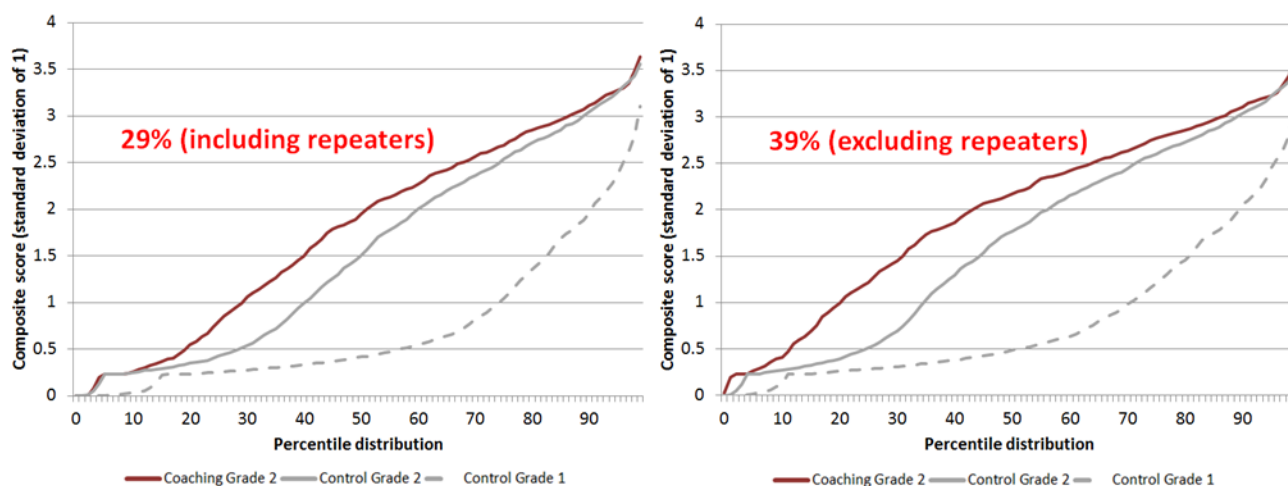
The graphs immediately make clear that the estimated impact expressed relative to a year of learning depends largely on the point in the performance distribution one chooses. This is complicated by the presence of floor effects. For example, the impact on paragraph reading (Oral Reading Fluency) at the 25th percentile would be zero, since even in the intervention group, there were more than 25% of learners who could not read a single word. Yet at the 30th percentile the impact would now be infinitely large since those in the coaching group at this point in the rank

distribution were able to read 5 words, but in the control group more than 30% of learners could not read a single word.

In light of these interpretation challenges, perhaps the best approach is to calculate the entire area between the control and intervention lines for grade 2 and express this as a percentage of the entire area between the Control at grade 1 and the Control at grade 2. Using this approach, we estimate that the impact of coaching on letter recognition was 24% of a year of learning; the impact on word recognition was 28% of a year of learning; the impact on non-word recognition (decoding) was 35% of a year of learning; and the impact on Oral Reading Fluency was 32% of a year of learning.

An additional approach adopted to estimate impact sizes relative to a year of learning, was to pool the data for Waves 2 and 3, keep only the common items, and then re-run the Principal Components Analysis to obtain composite scores that are comparable across Waves 2 and 3. For this approach, two writing items were also included which were common across the two waves. After deriving these composite scores, percentile plots were produced, yielding the “year of learning” graphs presented in Figure 9 below. After calculating the ratio of the area between the Coaching group and the control group at the end of Year 2 relative to the area between the control group at the end of Year 1 and the control group at the end of Year 2, the result was that the estimated impact of being in the Coaching intervention was 29% of a year of learning. If we exclude those repeating grade 1 from the calculation (since they were only exposed to the intervention in Year 1), the estimated impact is found to be 39% of a year of learning for those who were exposed to two years of the intervention.

Figure 9: “Year of learning” graphs based on a composite score derived from common items across Waves 2 and 3



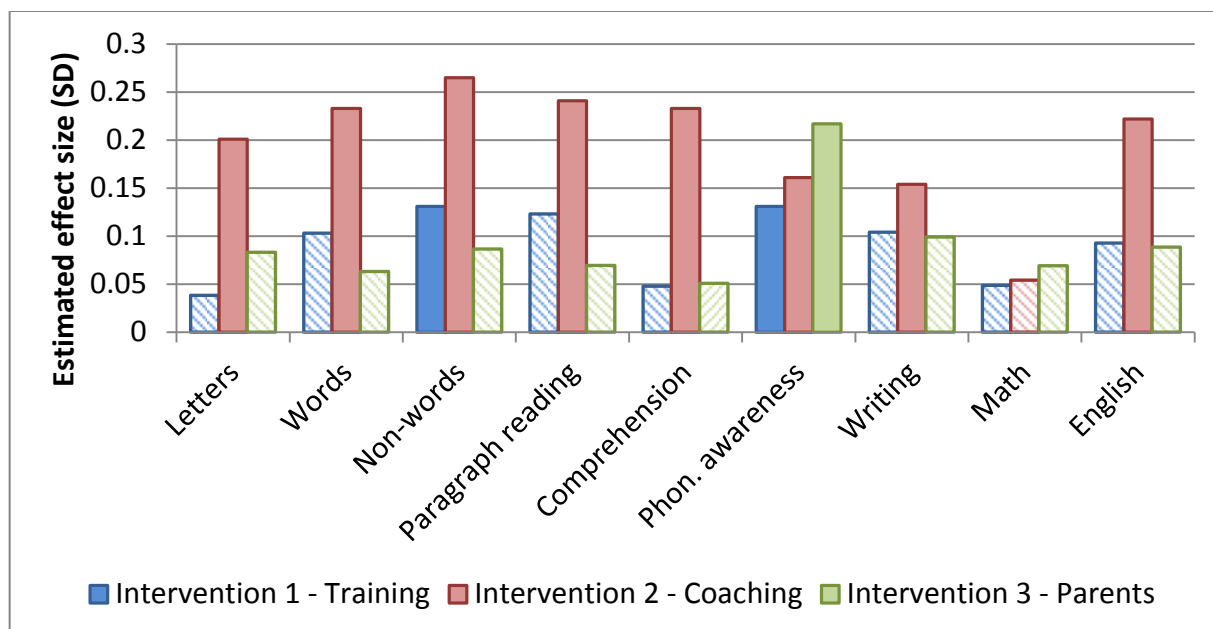
Impacts on various components of learning

We measure the impact of the Home Language Literacy interventions on letter recognition, word recognition, non-word recognition, paragraph reading (oral reading fluency), phonological awareness, comprehension, writing and two additional school subjects, English and mathematics – in case there were spillover effects. Figure 10 presents the results of the analysis of the impact on reading skills, with effect sizes expressed in terms of standard deviations. Fortunately, there were no negative effects of any intervention on any sub-test.

Although the Training intervention had moderate positive effects on some of the sub-tests, the Coaching intervention registered statistically significant positive effects on all Home Language Literacy measures, with similar effect sizes across the sub-tests. There was no significant effect of the Coaching intervention on the short mathematics test that was administered. This means that we have no evidence of a negative effect through crowding out of teaching time for mathematics. Interestingly, we observe a significant positive effect on English. This might be attributable to an improved underlying language ability (obtained through the home language intervention) or simply due to improved classroom management and transferable instructional methods acquired by the teacher through the Coaching intervention. Either way, this is an encouraging finding for the Coaching intervention.

Although the overall impact of the parent intervention was small, it does appear to have had a significant positive impact on phonological awareness. This was probably the specific reading skill that was most directly targeted through the parent meetings. Sound games were a key method taught to parents to use at home in the development of their child's phonological awareness.

Figure 10: Effects on sub-tests



Note: Solid bars represent statistical significance at the 90% confidence level.

Who benefits most from the interventions?

Boys catch up to some extent: The effective Coaching intervention is helping boys catch up some of the way to girls. Although girls still perform better than boys in the Coaching group, the gap is smaller than it is in the Control group.

Impact concentrated in urban schools: For all three interventions, the observed impacts are larger in urban township settings, but there is no measurable impact in deep rural settings. This means that we may need to approach interventions in rural schools differently. Table 8 indicates that the interaction between treatment and “rural” is negative and significant for all three intervention groups. Each intervention had a large estimated impact in urban township settings.

Table 8: Estimated treatment effects for urban and rural schools

| | Rural interaction |
|-------------------------|----------------------|
| Intervention 1 | 0.450*** (0.165) |
| Intervention 2 | 0.753*** (0.149) |
| Intervention 3 | 0.386*** (0.135) |
| Rural_X_ Intervention 1 | -0.391** (0.182) |
| Rural_X_ Intervention 2 | -0.631*** (0.177) |
| Rural_X_ Intervention 3 | -0.312* (0.162) |
| Rural | 0.316** (0.139) |
| Constant | -2.081*** (0.449) |
| Observations | 3,781 |
| R-squared | 0.181 |

Notes: Standard set of controls included
Cluster robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

A critical question is why were the interventions unsuccessful in rural settings? A number of things can be said to describe the differences between urban and rural settings. Rural schools are in poorer communities on average. Rural schools are smaller on average. Some small rural schools have multi-grade classes and this is not compatible with the grade-specific EGRS lesson plans. Class size is much the same on average. Both learner and teacher absenteeism is higher in rural areas. Parents are less educated in rural areas. However, none of these factors on their own seem to account for the urban-rural treatment heterogeneity. It is possible that a combination of these factors together make rural settings less conducive to interventions such as those run in EGRS.

The implementing agent, Class Act, also provided information about factors that caused disruptions to schooling and to the interventions, and the information was disaggregated by urban-rural. Table 9 describes the differences. On all observed dimensions teaching time was much more frequently lost in rural schools compared to urban schools. Poor weather, social unrest, teacher absenteeism and memorial services were especially frequent obstacles to normal schooling.

Table 9: Factors causing disruptions to schooling and EGRS interventions

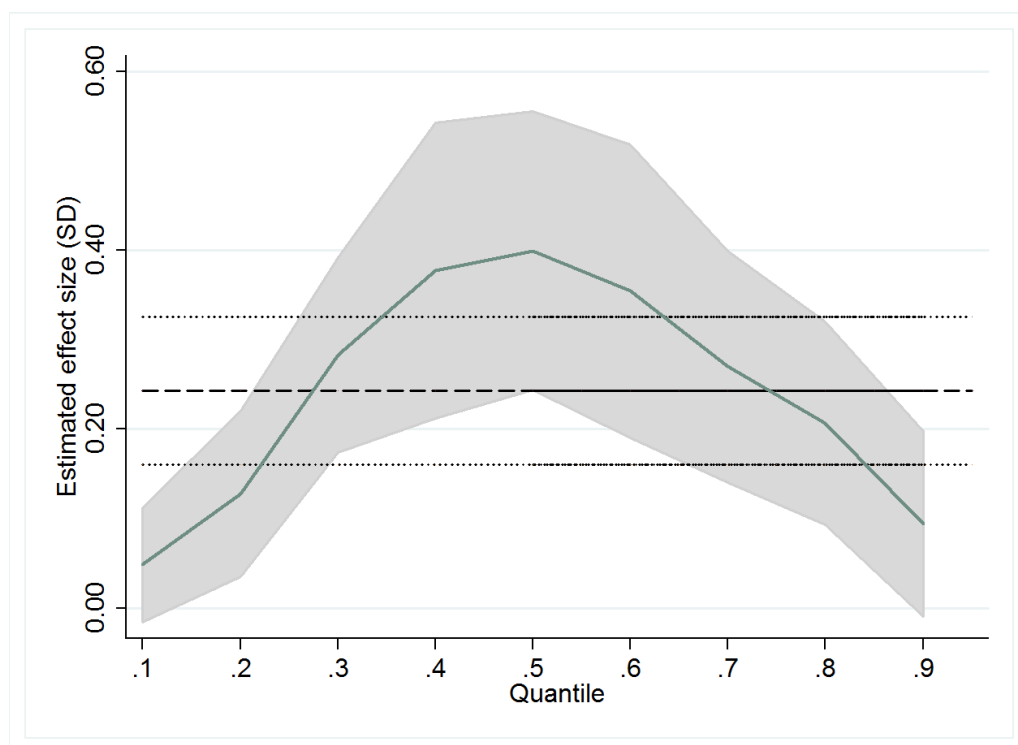
| | URBAN | RURAL |
|-----------------------------------|-------|-------|
| School access in rainy season | 0% | 48% |
| Disruptions as a result of unrest | 7% | 39% |
| Teacher absenteeism | 7% | 35% |
| Memorial services | 0% | 33% |
| Choir competitions | 14% | 21% |
| Sporting events | 7% | 24% |
| Poor cell phone coverage | 0% | 42% |
| OTHER (e.g. No electricity) | 7% | 21% |

Source: Class Act monitoring data

Large-classes benefited most: Both the teacher support interventions (“Training” and “Coaching”) had the largest impacts in relatively large classes (38 to 45 learners). In smaller classes, it may be that teachers in the control schools are already able to effectively manage classrooms, provide structured learning and differentiated attention to a variety of learners. However, in larger classes the EGRS interventions helped teachers to provide better instruction in a challenging environment. Both of the pedagogic interventions emphasised good classroom management practices such as how to reorganise classrooms, work in small groups while keeping the larger classroom occupied and bring routines and predictability to the classroom. However, in the very largest classes (more than 50 learners), the impact of the EGRS interventions were smaller, possibly indicating that beyond a certain threshold it remains difficult to conduct effective teaching. This emphasises the need to eliminate excessive class sizes (50+) in the Foundation Phase.

Middle-to-top performing learners benefited most: The impact of the Coaching intervention is largest for children in the middle and upper part of the achievement distribution with small or negligible impacts for the weakest performing children. The method we use to measure this is known as quantile regression. This estimates the effect of the intervention at various points in the distribution of the performance outcome. It asks, for example, what is the impact on the 10th percentile of performance, on the 20th percentile, on the 30th percentile, etc. Figure 11 presents the results of quantile regression measuring the effect of the Coaching intervention on the Year 2 composite score. The line plots the estimated effects across the performance distribution, while the shaded area represents the 95% confidence interval around the estimated effects. The impact on test scores was near zero at the 10th percentile of the distribution, and then quickly rose across the distribution, peaking at the 50th percentile. Importantly, there is no evidence of a negative effect for any part of the performance distribution. One implication of this finding is that structured pedagogic programmes that make use of lesson plans may benefit certain groups of children more, depending on the level at which the lessons are set.

Figure 11: Quantile regression of Coaching Intervention impact on Year 2 scores



How much did teaching practice and parent behaviour shift in response to EGRS interventions?

Through the use of mixed methods research (teacher questionnaires in all 230 schools, lesson observations in 60 schools and a set of detailed case studies), we investigate underlying change mechanisms by observing how the learning environment, teaching practice, and classroom activities changed as a result of the programmes.

An important preliminary point: Teachers in the Coaching schools were considerably more likely to report feeling a high level of professional support than those in the control schools, with teachers in the Training group also more likely to experience high professional support to some extent.

Table 10: Teachers' experiences of professional support

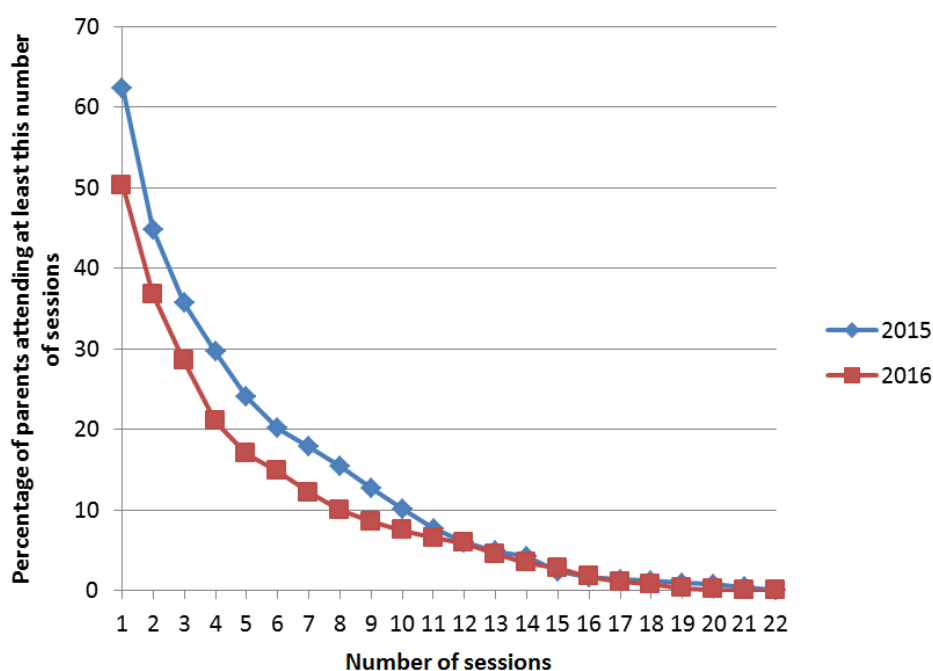
| | Control | Training | Coaching | Parents |
|-----------------------------------------------------------------------------|---------|----------|----------|---------|
| "I feel supported and recognised for my work" | 53% | 62% | 82% | 49% |
| "I regularly meet with people who provide mentoring and curriculum support" | 52% | 57% | 84% | 45% |

Two other results are worth emphasising. First, even though there are no large differences in access to graded readers, the lesson observations reveal that **far more learners are actually reading graded readers in the Coaching and Training schools. This increase is substantially larger for teachers who received Coaching relative to teachers who received Training.** Second, even though we find no change in the probability *that* learners practice reading in the classroom, there is a noticeable difference in *how* they practice reading: **Teachers in both Training and Coaching groups are more likely to do group-guided reading, resulting in more**

opportunities for learners to receive individual attention. The impact is, again, larger for teachers who received Coaching relative to Training. These results suggest that there are some teaching practices such as group-guided reading that are difficult to enact and require additional development to be effective. They also reveal an important interaction between resources and teaching practice: **graded readers are only useful if teachers have developed the skills to use them effectively in the classroom.**

Low attendance was a major limitation in the Parent intervention, as illustrated in Figure 12. In both 2015 and 2016, large proportions of parents did not attend any meetings. In 2015, just over a third of parents attended at least three sessions while in 2016 just under a third attended at least three sessions. Nevertheless, compared to the control group, parents in this intervention group reported attending a significantly higher number of parent meetings at their school on average. However, no other indicators of parental involvement in home reading or educational activities shifted substantially, confirming that there was no large change in parental behaviour in response to the intervention.

Figure 12: Distribution of total parent meetings attended, 2015 and 2016



Source: Class Act parent attendance records

A deeper analysis of classroom change: Summary of the 60-school lesson observation study

The main strength of Randomised Control Trials (RCTs) is their internal validity in measuring the causal impact of particular programmes. In other words, if outcomes end up higher in a group that received an intervention, we know that this is *because* of the intervention and we can make a quantitative estimate of that impact. But in order to gain a deeper understanding of *why* and *how* a programme may or may not have achieved its desired outcomes, one needs to complement the quantitative estimates of causal impact with mixed methods research. To better understand which

mechanisms were affecting the change in Interventions 1 and 2, a classroom observation study was commissioned.

The study was conducted in 60 of the schools that participated in the EGRS. A stratified random sample of 20 schools from each of the Control, Intervention 1 and Intervention 2 groups was chosen to form part of the study. In each of the schools, three different types of evidence were collected: (1) lesson observations; (2) evidence of work done in learners' workbooks and exercise books, as well as the review of various teaching documents and; (3) information from the teacher based on an interview.

Comparing the three different groups of schools, it emerged that the intervention schools were performing notably better than the control schools in the following themes: 'Teaching and Learning Environment'; 'Planning and Curriculum Coverage' and 'Classroom Management'. The main differences in the 'Teaching and Learning Environment' were the increased availability of display material (for example flashcards), a classroom arrangement that is more conducive to reading, and increased availability of reading books in the intervention classrooms.

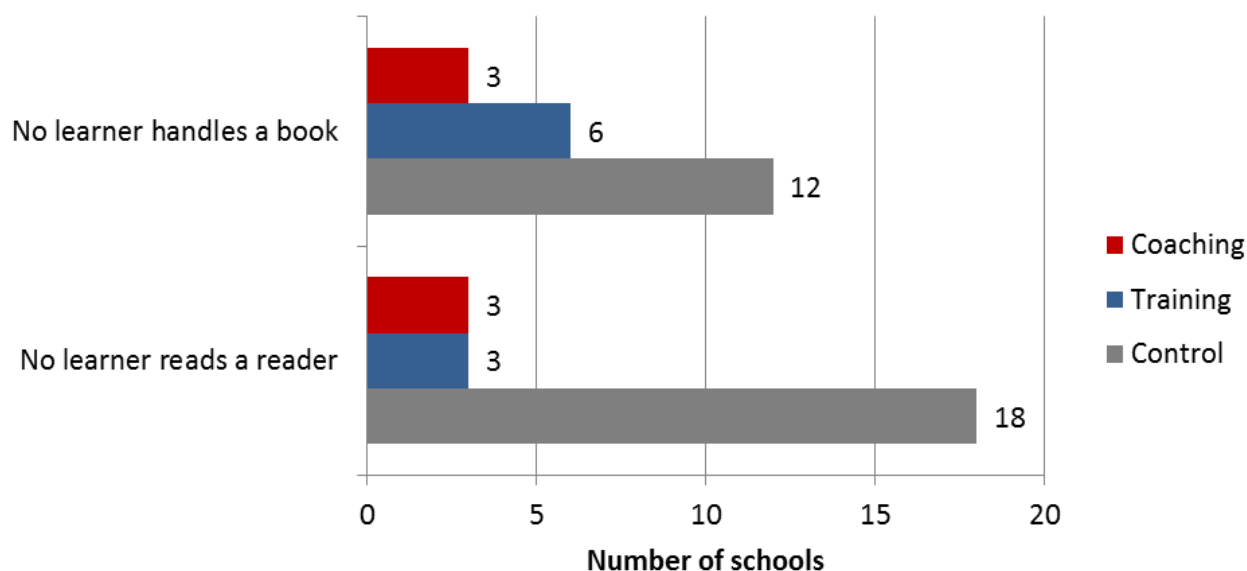
The scripted lesson plans provided through the programme proved to be hugely beneficial in translating the curriculum into daily lessons with detailed activities, which in turn improved 'Planning and Curriculum Coverage'. The specificity of the EGRS lesson plans was visibly different from the lesson plans used by the Control group's teachers and included important aspects such as vocabulary development. The benefit of greater specificity is especially clear with regards to vocabulary development, where teachers in Intervention 1 and 2 schools were much more likely to engage the learners in vocabulary development during the observed lesson. The EGRS lesson plans also provided teachers with a more accurate understanding of the size and scope of the curriculum that needs to be covered across the year, and provided them with a mechanism for tracking their own progress. The teachers in Intervention 2, however, were more likely to actually track their own progress and to be up to date in covering the curriculum. Evidence of increased curriculum coverage in Intervention 2 schools was found in the lessons observed, as well as in the learners' workbooks. The increased curriculum coverage meant that learners were more often engaged in writing activities and therefore learners in the intervention schools were less often observed being uninvolved in class. Although teachers in the intervention schools were observed to have a more realistic understanding of the curriculum scope, they still did not necessarily have a sufficient understanding of the cognitive demand required by the curriculum.

With regards to classroom management it was found that in 90% of the Intervention 2 classrooms no time was lost due to learners not being involved, whereas this was the case in 75% of the Intervention 1 classrooms. The evidence of more writing exercises in the learners' workbooks in the Intervention 1 and 2 schools corroborates the finding of learners being more involved and suggests that the improved classroom management is leading to increased curriculum coverage. In only 55% of the Control classrooms was no time lost due to learners not being involved.

In relation to the themes 'Opportunities to Write' and 'Use of Learning and Teaching Support Material' there were notable differences between Intervention 1 and Intervention 2 schools. Differences in the 'Use of Learning and Teaching Material' can be largely attributed to the prevalence of learners using storybooks and readers in class, as well as to the use of resources such as flashcards and charts by teachers during lesson observations in the Intervention 2 schools.

In 90% of the Control schools not a single learner was observed reading a graded reader, whereas this was commonly observed in the intervention classrooms (see Figure 13). These findings suggest that the EGRS interventions have been successful, not only in providing classrooms with the necessary readers, but also in ensuring that teachers make effective use of these resources.

Figure 13: Use of Learning and Teaching Support Material during lessons



With regards to ‘Opportunities to Write’, learners in Intervention 2 schools completed more writing exercises on average, specifically exercises pertaining to writing letters, short sentences and extended texts. Learners in Intervention 2 classrooms were also engaged in a wider variety of writing exercises overall were more likely to have their personally created dictionaries and to do more cursive writing exercises than learners in Intervention 1 classrooms.

Group guided reading provides a valuable opportunity for individualised and small group attention and was observed to occur more often in intervention classrooms. From evidence in the lesson observations it appears that in the majority of Control classrooms, the teachers’ in-class reading and phonics assessment was based on the class as a whole, rather than on individual learner proficiency.

The evidence found through the Lesson Observation Study suggests that the reading coaches played a critical role with regard to two aspects: (1) providing teachers with a more in-depth understanding of the enactment of the methodologies they were taught during the training; and perhaps more importantly, (2) supporting and motivating teachers to persist with the implementation of the programme. As mentioned above, there is significant evidence that Intervention 2 teachers were implementing the lesson plans as intended. Intervention 2 teachers were also more frequently seen providing different levels of readers to different ability groups in the lessons observed; doing a wider variety of writing activities during the Home Language lessons; covering the required pages in the DBE workbooks and covering more challenging aspects of the Grade 2 writing curriculum, especially writing sentences and extended texts.

Although Intervention 1 brought about significant changes in teachers' instructional practices, it seems that the reading coach component of Intervention 2 was the essential ingredient to encourage persistence in the curriculum-aligned learning programme. Available evidence therefore suggests that the 'triple cocktail' of lesson plans, high quality materials and coaching is necessary to affect real change in teachers' instructional practices.

Summary of the case studies

A set of case studies was undertaken by Dr Cheryl Reeves in four schools – two Training and two Coaching schools. Each case study involved lesson observations, teacher interviews and document reviews. A full report is available on these case studies.

A number of successful areas of the EGRS programmes were highlighted. Firstly, teachers were making daily use of the EGRS scripted lesson plans and regular use of the EGRS curriculum coverage trackers. Secondly, regular phonics, handwriting, group guided reading instruction and individual seatwork (writing) was taking place in EGRS classrooms. Thirdly, the provisioning of writing activities in the EGRS lesson plans was playing a role in motivating teachers to give classes more writing tasks, and learners were completing written work on most school days.

The case studies also identified several factors inhibiting programme impact. Particularly large classes made it difficult for teachers to provide learners with the individual attention they required. Secondly, there appeared to be an absence of a culture of reading for enjoyment and limited exposure of grade 2 learners to books besides the graded readers provided through EGRS and the DBE workbooks. Thirdly, teachers displayed a 'restricted' understanding of what it means to teach children to read independently – there was still an over-reliance on teacher-directed strategies (e.g. telling learners what words were).

A second set of case studies was conducted by Dr Kerry Dixon and Prof Brahm Fleisch in an additional four schools. These were all Coaching schools, selected at the extreme ends of the improvement spectrum based on the average performance on Wave 2 data. As in Dr Reeves' case studies, Dr Dixon and Prof Fleisch observed lessons, interviewed teachers, principals and other school staff, and reviewed classroom documents. A summary report is available on these case studies.

This report focuses on the complexities and nuances associated with the teachers' engagement with the various components and methods of the Coaching intervention. Although teachers lacked the vocabulary to talk about the five components of reading contained in the lesson plans, i.e. Phonological Awareness, Phonics, Vocabulary, Fluency and Comprehension (and writing), the strength of the lesson plans is that they incorporate all of these components in a set of standardised lessons, with simple, systematic routines. The lesson plans impacted both macro (across the academic year) and micro (within each lesson) pacing. Teachers singled out the positive types of learning that occurred during the coaching process, and signaled that a unique and helpful emotional environment was created by the coach. We also found that the new learning materials substantially contributed to improved instruction. The comprehensive set of 'word' flashcards were used extensively. Their popularity may be linked to teachers' familiarity with the 'look and say method' for teaching sight/high frequency words. The Vula Bula books were received very favourably by teachers and were observed in use. Teachers specifically noted that the books were

pitched at the correct level and were appropriately sequenced. A number of weaknesses were also observed. The phonics programme was not well understood by teachers. Group-guided reading, a key method for teaching reading was also not properly understood and was inadequately practiced. Whilst group-guided reading was essentially non-existent in Control schools (as evident in the 60-school lesson observation study), this indicates that even in the Coaching intervention there is a long way to go before reaching high quality instructional practice.

Cost-effectiveness Analysis

In thinking about which interventions are suitable to scale up, we need to consider both the impacts and costs of the programmes. There are various ways of doing a cost-effectiveness analysis. In most scenarios the Parent involvement programme is the most cost-effective (given its considerably lower cost) and Training is least cost-effective. However, since the overall impact of the Parents programme on the full evaluation sample is not statistically distinguishable from zero given conventional levels of significance, we believe it will be irresponsible to recommend scaling up the programme. Moreover, Coaching is more cost-effective in producing improvement in the comprehension test, which is arguably the most important goal of literacy. We therefore recommend wider implementation of the Coaching intervention, since it is the most cost-effective program that is known to work.

For estimates of costs, we use the programme budget for the third year of implementation for the Coaching and Training programmes. We chose the third year (2017), since this is at a point where a lot of the set-up challenges have been resolved and fixed costs have been paid (all the materials have already been developed). At this point in implementation, the largest cost-drivers in both programmes are variable costs (i.e. increase proportionally with number of teachers). One would therefore not expect the difference in per-learner costs to be much different when the programme is scaled up to more schools. Since the Parent programme was not implemented in the third year, we use the budget from the second year of implementation.

The total costs of implementation for the Coaching, Training and Parent interventions were R3 million, R2.34 million and R0.96 million respectively. Since these programmes were each implemented in 50 schools and the average number of grade 1 pupils in our sample of schools at the start of the programme was 74.6, the annual per-pupil costs are and R804, R626, and R256 respectively.

Given the impacts of 0.252, 0.12 and 0.1 SD increase for the respective programs, we can conclude that the Parent intervention was most cost-effective with a 0.38 SD increase for each R1,000 spent, compared to 0.31 and 0.18 SD increase for each R1,000 spent on the Coaching and Training interventions respectively.

Finally we consider a different performance metric: the increase in the number of pupils who pass the comprehension test. Here the Coaching intervention is most cost-effective. A learner is 12.3 percentage points more likely to pass the comprehension test per R1000 spent, compared to 6.6 and 3.3 percentage points in the Parents and Training arms respectively.

Conclusions and Recommendations

1. **Structured programmes with coaches help:** A structured learning programme aligned to the NCS, together with additional high quality reading support materials (graded reading books, flash cards, posters), can make a significant difference to learning outcomes, if accompanied by effective and carefully monitored support to teachers (coaches).
2. **Coaching is the best alternative:** Whereas previously very little evidence existed about effective large-scale teacher support modalities in South Africa, we now have evidence that on-site coaching to Foundation Phase teachers can shift learning outcomes, and that this is a cost-effective strategy. Modelling of lessons, in a safe space, as they navigate the lesson plans for teaching learners to read is critical.
3. **Direct in-service training is better than ‘train-the-trainer’ models:** Direct in-service training of teachers (4 two-day workshops over the course of 2 years), while less effective than on-site coaching, is in turn likely to have more impact than “cascade” models where specialists “train the trainers” who then interact with teachers.
4. **Existing subject advisers cannot fulfil the role of a coach:** The low ratio of subject advisors to schools (especially in the Foundation Phase) makes it impossible for subject advisors to fulfil the role of reading coaches, as implemented in EGRS; nor do we recommend increasing the number of subject advisors to allow this since the recruitment process, oversight structures and modus operandi of the coaches is different to that of subject advisors.
5. **Prioritising schools is a viable option:** On-site coaching interventions could be implemented in priority schools (e.g. 100 or 500 schools in a province) on a temporary basis (e.g. 2 years at a time) and through independent contracting and oversight structures. The cost for 100 schools would be R6 million at current prices.
6. **Develop reading norms in the African languages:** Reading norms cannot simply be adapted from one language into another due to differences in language structures. It is a complex exercise requiring longitudinal data. Therefore, the EGRS data could be used towards the development of reading norms in the African languages.
7. **Learning from EGRS:** Other large scale intervention initiatives such as those administered by the NECT could draw on the lessons of the EGRS and extend successful programmes to selected schools and districts.
8. **Parental involvement needs further research & may be promising:** Whilst parental involvement is a hugely deterministic factor in a child’s learning outcomes, the biggest challenge from a policy perspective is how to shift parent involvement at scale. Given the potential cost-effectiveness of such interventions, researchers and policy-makers should continue to investigate mechanisms to do so.
9. **Learning what works in deep rural settings:** Formative research and subsequent impact evaluation is required to figure out what types of school support programmes make a meaningful difference in deep rural settings.

The partners

An organisation called “Class Act” has been appointed to run the three interventions on behalf of the DBE for the purposes of this impact evaluation. Programme interventions are being funded by a coalition of donors, including the ZENEX Foundation, UNICEF, Anglo American and the Department of Planning, Monitoring and Evaluation in the Presidency. The evaluation side of the project is being supervised by the Research Team while the data collection and capturing is being managed by South Africa’s Human Sciences Research Council (HSRC) who also works closely with the Research Team on instrument development. The evaluation is being funded by the International Initiative for Impact Evaluation (3ie).

Bibliography

- Adukia, A. (2017). Sanitation and Education. *American Economic Journal: Applied Economics*, 9(2), 23-59.
- Altman, E. (2013). *Childhood nutrition in Chile: An assessment of a national school food program*. Atlanta: Thesis submitted to the Emory College of Arts and Sciences of Emory University in partial fulfilment of the requirements of the degree of Bachelor of Arts with Honors.
- Andrabi, T., Das, J., & Khwaja, A. (2015). Report cards: The impact of providing schools and child test scores on educational markets. Unpublished.
- Banerjee, A., Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of Participatory Programs: Evidence from a Randomised Evaluation in Education in India. *American Economic Journal: Economic Policy*, 2(1), 1-30.
- Conn, K. (2017). Identifying effective education interventions in Sub-Saharan Africa: A meta-analysis of rigorous impact evaluations. *Review of Educational Research*.
- Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidhara, K., & Sundararaman, V. (2011). School inputs, household substitution and test scores. *National Bureau of Economic Research Working Paper* 16830.
- Diagne, A., Lô, M., Sokhna, O., & Diallo, F. (2014). Evaluation of the impact of school canteen programmes on internal efficiency of schools, cognitive acquisitions and learning capacities of students in rural primary schools in Senegal. . *Partnership for Economic Policy working paper*.
- Ebenezer, R., Gunawardena, K., Kumarendran, B., Pathmeswaran, A., Jukes, M., Drake, L., et al. (2013). Cluster-randomised trial of the impact of school-based deworming and iron supplementation on the cognitive abilities of school children in Sri Lanka's plantation sector. . *Tropical Medicine and International Health*, 18, 942-951.
- Fleisch, B., Schöer, V., Roberts, G., & Thornton, A. (2016). System-wide improvement of early-grade mathematics: New evidence from the Gauteng Primary Language and Mathematics Strategy. *International Journal of Educational Development*, 49, 157-174.

- Freeman, M., Greene, L., Dreibelbis, R., Saboori, S., Muga, R., Brumback, B., et al. (2011). Assessing the impact of a school-based water treatment, hygiene and sanitation programme on pupil absence in Nyanza Province, Kenya: a cluster-randomized trial. *Tropical Medicine and International Health*, 17(3), 380-391.
- Glewwe, P., Hanushek, E., Humpage, S., & Ravina, R. (2014). School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010. In P. Glewwe (Ed.), *Educational Policy in Developing Countries*. Chicago and London: University of Chicago Press.
- Glewwe, P., Illias, N., & Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics*, 2(3), 205-227.
- Glewwe, P., Kremer, M., & Moulin, S. (2009). Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, 1(1), 111-135.
- Glewwe, P., Kremer, M., Moulin, S., & Zitzewitz. (2004). Retrospective vs. prospective analysis of school inputs: the case of flip charts in Kenya. *Journal of Development Economics*, 74, 251-268.
- Hanushek, E. (2010). The economic value of higher teacher quality. *National Center for Analysis of Longitudinal Data in Education Research, Working Paper No.56*.
- Ismail, S., Jarvis, E., & Borja-Vega, C. (2012). *Guyana's hinterland community-based school feeding program (SFP)*. Washington, DC: World Bank.
- Jukes, M., Zuilkowski, S., Parawan, A., & Lee, S. (2014). Evaluating the cognitive impacts of school health in the Philippines. *International Journal of Child Health and Human Development*, 7, 55-66.
- Kremmer, M., Brannen, C., & Glennerster, R. (2013). The challenge of education and learning in the developing world. *Science*, 340(6130), 297-300.
- Krishnaratne, S., White, H., & Carpenter, E. (2013). *Quality education for all children? What works in education in developing countries*. New Delhi: International Initiative for Impact Evaluation.
- Luo, R., Shi, Y., Zhang, L., Liu, C., Rozelle, S., Sharbono, B., et al. (2012). Nutrition and educational performance in rural China's elementary schools: Results of a randomised control trial in Shaanxi province. *Economic Development and Cultural Change*, 60, 735-772.
- McEwan, P. (2012). Improving learning in primary schools of developing countries: A meta-analysis of randomised experiments. *Review of Educational Research*, 20(10), 1-42.
- McEwan, P. (2013). The impact of Chile's school feeding program on education outcomes. *Economics of Education Review*, 32, 122-139.
- Miguel, E., & Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 82, 159-217.
- Murhaldiharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of political economy*, 119(1), 39-77.

- Murnane, R., & Ganimian, A. (2014). Improving educational outcomes in developing countries: Lessons from rigorous evaluations. *NBER Working Paper No 20284, National Bureau of Economic Research, Cambridge, MA.*
- Nguyen, T., & Lassibille, G. (2008). Improving management in education: Evidence from a randomised experiment in Madagascar. Unpublished manuscript 6.
- Nonoyama-Tarumi, Y., & Bredenberg, K. (2009). Impact of school readiness program interventions on children's learning in Cambodia. *International Journal of Educational Development, 29*(1), 39-45.
- Pandey, P., & Goyal, S. a. (2011). Does information improve school accountability? Results of a large randomised trial. Washington, DC: World Bank.
- Piper, B. (2009). *Integrated Education Program: Impact Study of SMRS Using Early Grade Reading Assessment in Three Provinces in South Africa*. Pretoria: United States Agency for International Development.
- Piper, B., Zuilkowski, S., & Mugenda, A. (2014). Improving reading outcomes in Kenya: First-year effects of the PRIMR initiative. *International Journal of Educational Development, 32*7, 11-21.
- Piper, B., Zuilkowski, S., & Mugenda, A. (2014). Improving reading outcomes in Kenya: First-year effects of the PRIMR Initiative. *International Journal of Education Development, 37*(11-21), 11-21.
- Popova, A., Evans, D., & Arancibia, V. (2016). Inside In-Service Teacher Training: What works and how do we measure it? Washington, DC.: Rise Programme.
- Sabarwal, S., Evans, D., & Marshak, A. (2014). The Permanent Input Hypothesis: The Case of Textbooks and (No) Student Learning in Sierra Leone. *World Bank Policy Research Working Paper 7021*.
- Sailors, M., Hoffman, J., Pearson, P., Beretvas, S., & Matthee, B. (2010). The Effects of First- and Second-Language Instruction in Rural South African Schools. *Bilingual Research Journal, 21*-41.
- Simeon, D., Grantham-McGregor, S., Callender, J., & Wong, M. (1995). Treatment of *Trichuris trichiura* infections improves growth, spelling scores and school attendance in some children. *Community and International Nutrition, 18*75-1883.
- Snilstveit, B., Stevenson, J., Menon, R., Philips, D., Gallagher, E., Geleen, M., et al. (2016). *The impact of education programmes on learning and school participation in low- and middle-income countries*. London: International Initiative for Impact Evaluation.
- Watkins, W., Cruz, J., & Pollitt, E. (1996). The effects of deworming on indicators of school performance in Guatemala. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 90*, 156-161.
- Wong, H., Shi, Y., Luo, R., Zhang, L., & Rozelle, S. (2014). Improving the health and education of elementary school children in rural China: Iron supplementation versus nutritional training for parents. *Journal of Development Studies, 50*, 502-519.



Education and Sport Development
 Department of Education and Sport Development
 Departement van Onderwys en Sportontwikkeling
 Lefapha la Thuto le Tlhabololo ya Metshameko
NORTH WEST PROVINCE



basic education
 Department:
 Basic Education
REPUBLIC OF SOUTH AFRICA



**planning, monitoring
& evaluation**
 Department:
 Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

