

The Second Early Grade Reading Study

Early Perspectives

The situation at the start of Grade 1.

Contents

Executive Summary.....	6
1. Introduction.....	8
2. Background.....	9
3. Study Design	10
3.1. Research Site	10
3.2. Interventions.....	12
3.3. Theory of Change.....	14
3.3.1. Teaching English as First Additional Language.....	15
3.3.2. Common aspects of the two interventions.....	15
3.3.3. Intervention 1 specifically:	16
3.3.4. Intervention 2 specifically:	16
3.4. Evaluation	16
3.4.1. Sample selection and intervention assignment	17
3.4.2. Instrument Design.....	17
3.4.3. Data Collection.....	18
3.5. Ethical Clearance.....	21
4. Baseline Results	21
4.1. Sample Characteristics	21
4.1.1. Learner Characteristics	22
4.1.2. Teacher Characteristics.....	24
4.1.3. School and Principal Characteristics.....	26
4.2. Learner Assessment.....	27
4.3. Balance Test.....	35
4.4. Which characteristics are correlated with learner test scores?	37
5. Sensitivity Checks.....	41
6. USAID indicators	43
7. Next Steps in the project.....	45

List of Figures

Figure 1: Map of the research site	11
Figure 2: Who is responsible for your child's reading?	24
Figure 3: How often do you speak English to your child?	24
Figure 4: Teacher age profile	24
Figure 5: Resources used for EFAL lesson planning	25
Figure 6: Two most important lesson types	25
Figure 7: Subject teachers enjoy teaching most	25
Figure 8: Teacher's use of technology	26
Figure 9: Principal age profile	26
Figure 10: Instructional Leadership - Time is dedicated to EFAL?	27
Figure 11: Instructional Leadership -When to start with EFAL?	27
Figure 12: Overall learner performance distribution	29
Figure 13: Histogram – Naming Animals.....	30
Figure 14: Histogram - Word recall	30
Figure 15: Histogram - Phoneme Isolation	32
Figure 16: Histogram - Story Comprehension	32
Figure 17: Distribution of learner performance on Letter Sound Recognition	33
Figure 18: Sub-task 8 - Visual perception	34
Figure 19: Composite learner score by intervention group.....	35
Figure 20: Sub-task performance by intervention group	36
Figure 21: Language Difference in Phoneme Isolation	39
Figure 22: Speaking English at home and learner performance	40
Figure 23: Reading at home and learner performance.....	40

List of Tables

Table 1: Percentage of learners tested during baseline data collection.....	19
Table 2: Percentage of teacher interviews conducted during baseline data collection	20
Table 3: Number of principals interviewed during baseline data collection	20
Table 4: Percentage of Home Background Questionnaires returned	20
Table 5: School Characteristics.....	22
Table 6: Learner characteristics	23
Table 7: Average performance per sub-task	28
Table 8: Summary Statistics per sub-task	28
Table 9: Subtask 1 - Naming Animals	30
Table 10: Sub-task 2.1 Word Recall.....	30
Table 11: Sub-task 2.2. Non-word Recall	31
Table 12: Sub-task 3 - Phoneme Isolation	32
Table 13: English Receptive and Expressive Vocabulary	35
Table 14: Balance tests per sub-task.....	36
Table 15: Balance test on Teacher English Proficiency	37
Table 16: Learner Characteristics correlated with baseline performance.....	38
Table 17: Learner Characteristics correlated with learner performance per sub-task.....	39
Table 18: Learner background factors correlated with baseline performance	40
Table 19: Community level factors	41
Table 20: Robustness check - Re-allocating intervention status	42
Table 21: Specific USAID Indicators.....	43
Table 22: Custom USAID Indicators	44

Acronyms

CAPS	Curriculum and Assessment Policy Statement
DBE	Department of Basic Education
EFAL	English as First Additional Language
EGRA	Early Grade Reading Assessment
EGRS	Early Grade Reading Study
EGRS II	Second Early Grade Reading Study
HOD	Head of Department
ICT	Information Communication Technology
LoLT	Language of Learning and Teaching
LTSM	Learning and Teaching Support Material
NECT	National Education Collaboration Trust
OLS	Ordinary Least Squares
PIRLS	Progress in International Reading and Literacy Study
RCT	Randomised Control Trial
SADTU	South African Democratic Teacher Union
SMT	School Management Team
USAID	United States Agency for International Development

Executive Summary

Early literacy has been taking centre stage as one of the most important aspects in ensuring sustained learning throughout an individual's schooling career. Literacy allows individuals to access information, improve their knowledge, increase their productivity and contribute to society using their full potential. Literacy therefore does not only hold major benefits for the individual, but has spill-over effects to the broader society, and at a national level leads to economic growth. Ensuring that all individuals are literate by the end of Grade 12 is therefore of critical importance to the South African Department of Basic Education (DBE).

In pursuing this objective, the DBE in collaboration with the University of Witwatersrand, has embarked on the Early Grade Reading Study (EGRS) series in 2015. Both the first and second EGRS aim to evaluate and trial more effective ways of supporting and training teachers in order to influence both the quality of instructional practice and the pace of teaching. The theory of change is therefore that improvement in these two aspects of instructional practice will lead to more effective teaching, and will subsequently lead to improved learner performance. EGRS I supported teachers in the teaching of Setswana as Home Language in the Foundation Phase, whereas EGRS II will be supporting teachers in the teaching of English as First Additional Language in the Foundation Phase.

Building on the experience of EGRS I, EGRS II evaluates two different interventions which support teachers to teach more effectively. In both interventions teachers are provided with lesson plans and additional reading resources, as well as instructional coaching and training. Two different strategies to training and coaching are used: (i) the traditional face-to-face block training, complemented with regular on-site coaching, and (ii) the traditional face-to-face block training, combined with an ICT component that includes tablets provided to teachers and a virtual coach which provides on-going support through cell phone messaging and phone calls to the teacher.

EGRS II is designed as a Randomised Control Trial (RCT) and schools are therefore randomly assigned to each of the intervention groups (50 schools each), as well as the control group. (80 schools) This ensures comparability between the groups, with the only difference between them being the treatment allocation. The random assignment is fundamental in the design of the study and will serve as the basis of making claims about the causal impact of each of the intervention groups on the English reading outcomes of learners.

The interventions of the study will be implemented in Mpumalanga in Grade 1 in 2017 and Grade 2 in 2018. To evaluate the impact of the interventions, Grade 1 learners was tested at the start of Grade 1 in 2017 and will be tracked throughout the course of the study, participating in assessments at the end of Grade 1 (2017) and at the end of Grade 2 (2018).

The baseline data collection was administered at the end of January and start of February 2017 in all 180 schools. A random sample of 20 Grade 1 learners per school were selected to participate in the oral assessment of reading and pre-reading skills. The learner assessment was administered in the Language of Learning and Teaching (LoLT) of the school, which was

either isiZulu or SiSwati. Questionnaires were also administered to the school principal, to all Grade 1 teachers and to the parents of the 20 tested learners. The learner tests were adapted from the Early Grade Reading Assessment (EGRA) tool in isiZulu and SiSwati. The assessment covered the following skills: Expressive vocabulary, word recall and non-word recall, phoneme isolation, listening comprehension and visual perception. A task assessing learners' English vocabulary was also included.

The baseline testing confirms the success of randomisation, with no significant difference in the performance of the learners in the different treatment groups on the various sub-tasks of the test. Some of the sub-task did show some ceiling effects, whereas others showed some floor effects, but overall the test seems to provide good discrimination between learners of varying ability levels. The correlation between the sub-tasks are relatively low, which does seem to suggest that there is some noise in the information generated from the assessment.

The Grade 1 teachers in the sample are all female teachers. Just over a quarter of the teachers are older than 55 years old, whereas just over half of teachers are between the ages of 45 and 54 years old. This could potentially have implications for the implementation of the interventions, specifically since one of the interventions involve the use of technology. Given the South African context, however, this situation was expected and the interventions were designed with this in mind.

Interventions continued throughout 2017, with the initial training having been conducted just after the conclusion of the baseline data collection. Teachers further attended clustered workshop training for one day at the start of each term, and the reading coaches provided on-going support throughout the year. The midline data collection is taking place from 23 October to 10 November 2017, and will provide the data necessary to determine the impact of the interventions after one year of implementation. The interventions are set to continue to Grade 2 teachers in 2018, and the same learners will be tested at the end of Grade 2 in 2018.

1. Introduction

Reading is arguably the most important skill which a child needs to acquire during the foundation phase. If a child has not yet acquired the skill to read for meaning by the end of the third grade, all further learning will be constrained. This has implications on an individual level, as the skill of reading is a necessary requirement for the acquisition of all further learning, and subsequently leads to higher wages and a higher likelihood of employment. Moreover, higher literacy rates have societal benefits and are associated with higher economic growth, healthier societies and less crime.

In South Africa literacy rates are very low in all languages. Large scale tests of reading proficiency show that the majority of learners in South Africa cannot read for meaning by the end of Grade 4. This trend is worse when looking at only African language readers. Although, the South African language policy promotes the teaching of reading in learners' home language during the first three years of schooling, all further learning from the fourth grade will most likely be taking place in English for the majority of learners. This means that learners do not only need to be able to read for meaning in their Home Language, but also in English. English, however, is not only an additional language for the learners, but is also the additional language for the majority of teachers in South Africa. Teaching English as First Additional Language (EFAL) is therefore often done very scantily, with teachers lacking confidence in speaking English themselves. To address these systemic challenges, the Early Grade Reading Study II (EGRS II) aims to support and strengthen the teaching of early grade reading in English by providing teachers with structured lesson plans, additional learning and teaching support material, training sessions and reading coaches.

The Early Grade Reading Study II builds on the experience of the Early Grade Reading Study I that was conducted in the North West province. Both interventions aim to contribute to knowledge about how to support and train teachers to influence both the quality of instructional practice and the pace of their teaching. Improved instructional practices will enable learners to better understand and internalize the content of the lesson, whereas an increased pace will ensure that teachers cover the entire curriculum in the year. The theory of change is therefore that change in these two aspects of instructional practice will lead to more effective teaching, and will subsequently lead to improved learner performance.

For this reason the EGRS II interventions target both of these elements. To evaluate new ways in which to assist and support teachers to teach more effectively, teachers are provided with lesson plans and additional reading resources, as well as instructional coaching and training. Two different strategies to training and coaching are used: (i) the traditional face-to-face block training, complemented with regular on-site coaching, and (ii) the traditional face-to-face block training, combined with an ICT component that includes tablets provided to teachers and a virtual coach which provides on-going support through cell phone messaging and phone calls to the teacher. The lesson plans are fully aligned with the Curriculum and Assessment Policy Statement (CAPS) and give the teacher a sense of the break-down of the CAPS curriculum in daily lessons that will allow the coverage of the entire curriculum within the academic year. In

addition, the reading coaches encourage teachers to stay on track and teach each lesson, as well as support the teacher in teaching the sections which are found to be more challenging.

2. Background

Reading is often used as an indicator of how well an education system is delivering on its mandate. International assessments such as the Progress in International Reading Literacy Study (PIRLS) have shown that at Grade 4 and Grade 5 level, South African learners have not yet learned to read with meaning. In the 2011 prePIRLS assessment, 29% of learners did not even reach the low international benchmark and therefore, when reading information texts these learners could not make any inferences about logical connections to provide reasons, or interpret obvious reasons and causes when reading literary texts (Howie, et al., 2012). By these grade levels learners are required to use these specific skills to accumulate all further knowledge, signifying that for the largest majority of South African learners, further learning will always be constrained.

South Africa, like many other countries on the African continent, has a rich language heritage with eleven official languages. The language policy in South Africa explicitly promotes primary home language instruction in the foundation phase and the choice of the specific language of learning and teaching (LoLT) in the foundation phase is left to the governing bodies of schools. From Grade 4 onwards, learners are taught most subjects in either English or Afrikaans, but English is the language chosen by the majority of schools. Although African languages are not used as the LoLT from Grade 4 onwards, African language speakers are expected to continue studying their home language as a school subject until Grade 12. To ameliorate the language transition learners' face in Grade 4, English is introduced as an additional language from Grade 1. The language policy therefore promotes an additive approach to bilingualism where the home language is developed together with the additional language (Matjila & Pretorius, 2004).

Strengthening the teaching of English in the foundation phase is therefore critical to effect improved learning performance in the later grades. In understanding the main constraints to more effective teaching, school effectiveness studies in South Africa have found that primary school classrooms are mainly characterised by a lack of print material, a lack of opportunities for reading and writing, chorusing practices, low levels of cognitive demand and slow pacing (Taylor, 2007). However, while a range of factors influence learning outcomes, it is widely recognised that instruction or instructional practice is critical to improving learning outcomes (Coe, et al., 2014). One of the key characteristics of South African education is that the dualistic nature of learning outcomes between the wealthy and the poor, is mirrored by dual types of instructional practice happening in the schools serving these communities (Hoadley, 2012). It is likely that weak instructional practices have a causal impact on learning outcomes in the poorly performing part of the school system. To substantially shift achievement in the weak part of the schooling system it may be necessary to apply a comprehensive instructional change intervention, involving a set of coherent and aligned instructional inputs. For this study,

the instructional inputs include lesson plans, aligned learning materials and in-classroom support to teachers.

The study is designed as a randomised control trial (RCT) to allow the robust evaluation of the causal impact of the two interventions trialled. This impact evaluation will thus inform credible policy decisions around interventions that have been implemented on a relatively large scale. The main benefit of the RCT design is the inclusion of a randomly selected control group in the evaluation, which allows the measurement of the amount of learning that would have taken place in schools, should the interventions not have been implemented. This in turn provides a benchmark to which the learning gains in the intervention groups can be compared in order to establish the additional learning gained by implementing the interventions. Furthermore, the comparison of the two interventions with the control group also allows a direct comparison of the costs involved to effect the learning gains, and therefore allows the determination of the cost-effectiveness of each of the interventions.

The primary implementing partner is the South African government, in particular the Department of Basic Education. The national Department has partnered with the Provincial Education Department in Mpumalanga and the University of the Witwatersrand, to implement and evaluate the interventions in the province. Service providers have been appointed to assist with the implementation of the interventions, as well as the data collection for the evaluation of the programme. The evaluation side of the project is being supervised by the Research Team. The study is completely funded through USAID, with the funding housed and administered through the University of Witwatersrand.

3. Study Design

The EGRS II is designed as a Randomised Control Trial (RCT), which aims to measure the effects of each of the interventions and compare it to the situation among a comparison group of learners where typical teaching is taking place. The study entails the implementation of two early grade reading interventions in South Africa in 180 primary schools in the Mpumalanga Province during 2017 and 2018. The EGRS II focuses on the early learning of English as the First Additional Language and the interventions consist of lesson plans, additional reading resources and instructional coaching and training. Two alternative strategies to training and coaching are used: (i) the traditional face-to-face format, (ii) a combination of face-to-face training and an ICT component that includes tablets and cell phone messaging to the teacher. The sections to follow will elaborate on the research site, the specific interventions, the theory of change, the sample selection, the instrument design and the baseline data collection.

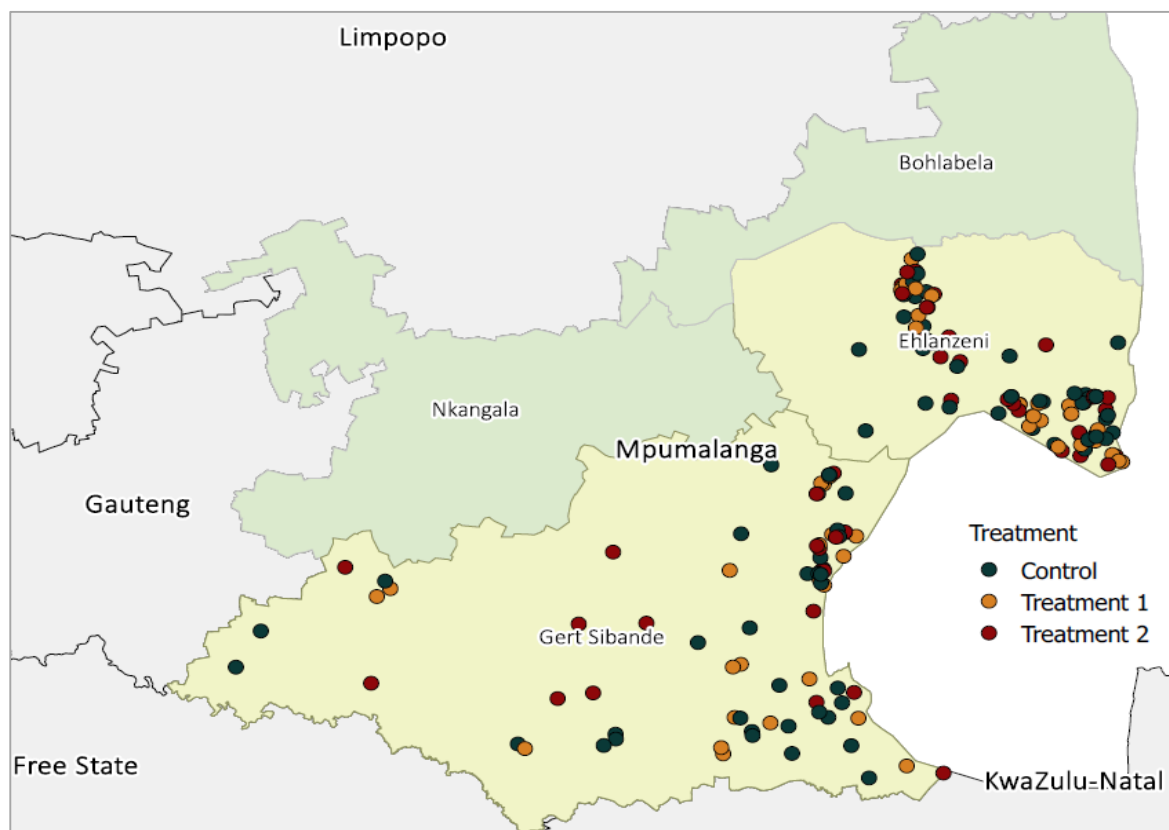
3.1. Research Site

In selecting Mpumalanga as the province in which to implement and evaluate the second Early Grade Reading Study, the following requirements were considered:

1. Close proximity to the Department of Basic Education in Pretoria to enable officials to monitor the implementation and evaluation at a relatively low cost.
2. A politically stable environment with minimal interruptions to the school year due to service delivery or other protests.
3. A relatively well-functioning Provincial Office which will be capable of providing the necessary support to the implementation of the intervention.
4. A relatively well-functioning school environment with limited teacher absenteeism.
5. Enough schools for an adequate sample, given the restrictions imposed.

A further consideration in selecting the districts in which to implement the study was the home languages spoken in each district. The interventions target English as First Additional Language, but the learner assessments contain sections assessing the learners' proficiency in their home language. These sections are particularly prominent in Grade 1 and Grade 2, and require fieldworkers to administer the assessments in the learners' home language. For this reason it is necessary to determine the extent of multilingualism in the schools of the respective districts. In the districts in Mpumalanga there is a large variety of home languages, but in the district of Ehlanzeni and Gert Sibande the majority of schools either have isiZulu or siSwati as the LoLT. The district of Bohlabela was excluded because the National Education Collaboration Trust (NECT) has already implemented lesson plans in the district.

Figure 1: Map of the research site



Relative to the other provinces in South Africa, Mpumalanga is a mid- to low- performing province. In the 2016 Matric Examination¹, Mpumalanga ranked fifth out of the nine provinces. Unfortunately, South Africa does not currently have a standardised assessment with which to compare schools, districts and provinces in the primary phase, but various other indicators can provide an indication of the current schooling conditions. The 2016 General Household Survey reports that 94.3% of 5 to 6 year olds in Mpumalanga are currently attending an educational institution and 91.4% of Grade 1 learners reported having attended Grade R prior to Grade 1. Poverty analysis shows that 28.4% of learners attending schools in Mpumalanga fall below the food poverty line (monthly per capita income is below R442.00) and 47% below the Lower Bound Poverty Line (monthly per capita income is below R660.00).

3.2. Interventions

The EGRS II focuses on the early learning of English as the First Additional Language by providing specific resources, training and on-going support to teachers. The interventions aim to support Foundation Phase teachers in their instruction of English as a First Additional Language, in accordance with the DBE's National Curriculum Statement, including the CAPS. Both interventions consist of three components: (1) lesson plans, (2) integrated learning and teaching support material and (3) instructional coaching and training. The key differences between the two interventions lies in the delivery mode of coaching support that teachers receive and the format in which teachers receive the lesson plans. In both interventions teachers receive initial block training, as well as one day training at the start of each school term. In the first intervention, the teachers receive a paper-based version of the lesson plans and benefits from regular on-site coaching. In Intervention 2, the teachers receive a tablet with an electronic version of the lesson plans, including various audio-visual resources, and are supported through an ICT coaching model that includes telephone calls and cell phone messaging to the teacher. The electronic lesson plans in the second intervention are delivered using an application which is specifically developed for the study, and the additional electronic resources include short training videos, sound clips of the phonics sounds, songs and rhymes and examples of learners' work. The focus of both the interventions is explicitly on how to deliver English First Additional Language instruction and on how to effectively use the new materials as well as other available materials - especially the government-provided workbooks - to most effectively teach reading in accordance with the National Curriculum Statement Grades R – 12.

Intervention 1 is implemented in 50 randomly selected schools and provides teachers with regular face-to-face coaching, as well as quarterly training workshops. The training workshops in both interventions are similar and focus specifically on supporting teachers with the integration of the lesson plans and LTSM into the teaching of EFAL, as well as on various methodologies for the teaching of reading and literacy in English. Unlike typical training

¹ The final school leaving examination in South Africa. This examination is standardised across all schools in South Africa, and is the only credible standardised assessment which can be used to make cross-province comparisons.

workshops, this is best understood as ‘just-in-time’ training: not on general principles, but specifically on how to implement the English FAL literacy programme in the upcoming weeks. The provision of specialist reading coaches is an alternative model of support, involving regular (once a month) in-school coaching. The reading coaches play a number of distinct and overlapping roles, firstly in modelling new practices and secondly in monitoring implementation fidelity. As the teachers become more familiar with the new practices, the coach’s role shifts from introducing new practices towards that of ‘critical friend’, observing and providing real-time input to teachers as they begin to master the new practice. The coaches also play an important monitoring role, providing regular feedback to the project management team on the level and quality of programme implementation. The monitoring is always done in an encouraging manner, which therefore contributes to teachers being more motivated and encouraged to implement the programme throughout the year. The reading coaches also initiate needs-driven clustered workshops throughout the school term with identified teachers invited to attend. Finally, teachers in both interventions are supplied with sets of LTSM including graded reading booklets, posters, big books, sentence strips and vocabulary words.

Intervention 2 is implemented in a different group of 50 randomly selected schools and uses a different model of teacher support and mode of delivering the lesson plan. The technology-supplemented intervention aims to provide new forms of support and guidance on teaching strategies through a range of resources, electronically available lesson plans and interactive support platforms that are available at all times to the teacher. The application further includes various resources such as video clips on best practice, and audio clips of the phonic sounds, the songs and the rhymes in the lesson plans. The intervention makes use of WhatsApp/SMS to create virtual reading coaching practices and virtual communities of practice. The virtual reading coach uses text messaging to communicate with teachers regularly, providing them with teaching tips on a weekly basis, and answering any questions they may have on the lessons that they have taught, or are preparing to teach the following day. The regular communication is important in building trust in the virtual reading coach as well as in the overall intervention. Secondly, the virtual reading coach also helps facilitate virtual teacher groups using an appropriate platform such as WhatsApp. This platform assists in building teacher networks where they can share their experiences and assist each other with useful tips and techniques.

	Intervention 1	Intervention 2
Provision of lesson plans	Paper based	Electronic On an application on a tablet
Provision of LTSM	Paper based: <ul style="list-style-type: none"> - Big books - Posters - Flash cards - Writing frames 	Paper based: <ul style="list-style-type: none"> - Big books - Posters - Flash cards - Writing frames

	Intervention 1	Intervention 2
Coaching	Coach visits teacher in her classroom. Once every three weeks.	Coach contacts teacher via telephone calls and instant messaging (Whatsapp). Once every two weeks.
Training	Initial training: 2 day block training Quarterly training: 1 day at the start of each term Needs-based training: As required	Initial training: 3 day block training Quarterly training: 1 day at the start of each term Needs-based training: None
Core methodologies	Paper based instructional manual with accompanying CD with phonics sounds	Application based instructions, Includes videos, sounds clips and photos of example writing

3.3. Theory of Change

The Early Grade Reading Study I showed that the combination of lesson plans, integrated LTSM and coaching was particularly important in affecting sustained behaviour change in the teaching of Home Language. In the Early Grade Reading Study II we are evaluating whether the same results can be obtained when applying the same programme in a different province and in a different subject (English First Additional Language). Furthermore, we are evaluating whether on-site, face-to-face coaching is essential, or technology can be utilised to reduce the costs of coaching with the same effects.

The Early Grade Reading Study II aims to effect behaviour change among teachers at a large scale, in line with both the curriculum and methodologies in which teachers were trained during the teacher training at the start of the programme. The lesson plans form the foundation of the interventions by integrating the different components of the programme. The lesson plans provide a focus for the coaching interactions and ensure the appropriate use of the learning and teaching support materials. The lesson plans therefore provide a mechanism to prompt the enactment of the behaviour change, whereas the coaching serves as an additional mechanism to encourage fidelity to the programme. The role of the learning materials is to provide the appropriate resources to ensure that learners are able to develop and consolidate knowledge and skills related to reading fluency, vocabulary development and shared reading.

Given the perceived high costs of direct in-class coaching, the study tests the relative cost-effectiveness of two kinds of capacity building: the traditional face-to-face model and the new combined face-to-face model supplemented by electronic support in the form of a tablet with video demonstration lessons and cell phone messaging support. The technology-supplemented intervention aims to provide new forms of support and guidance on teaching strategies through a range of materials, teaching guides, videos and interactive support

platforms that are available at all times to the teacher. While the Early Grade Reading Programme cannot address all the learning challenges, particularly for learners with severe learning disability, it aims to strengthen English reading performance for the majority of learners across the performance distribution.

3.3.1. Teaching English as First Additional Language

Both interventions are fully CAPS compliant and aim to support teachers in the enactment of the current South African curriculum for English as a First Additional Language. The focus of the study is to inform the system on ways to strengthen teaching and learning in EFAL in the foundation phase, in order to prepare learners for the language transition to English as the LoLT which happens in the majority of schools in Grade 4.

A range of factors influence learning outcomes, including school and non-school variables, but instruction or instructional practice is one major influence on learning. To substantially shift achievement in the weak part of the schooling system it may be necessary to apply a comprehensive instructional change intervention, involving a set of coherent and aligned instructional inputs. To address these complexities, the interventions provide teachers with instructional inputs such as lesson plans, aligned learning materials, and support and training to teachers.

3.3.2. Common aspects of the two interventions

The lesson plans provide specification of the new instructional practice including faster paced instruction, more appropriately sequenced content, and dramatically expanded pedagogic repertoires. In the primary school teaching of reading in English as the First Additional Language, the new expanded repertoires include systematic teaching of phonemic awareness and phonics, strategies that focus on increased reading speeds or fluency, shared reading strategies, vocabulary development and strategies that improve comprehension. The lessons require little additional lesson preparation from teachers, which makes the take-up of a more productive teaching practice more manageable. The faster paced instruction (relative to typical practice) also ensures that the teachers cover the full prescribed curriculum for the year. Furthermore, lesson plans free up teachers' time, because they no longer need to allocate as much time to planning. This could improve reading acquisition if teachers allocate this time to productive teaching activities.

The role of the learning materials is to provide the appropriate resources to ensure that learners are able to develop and consolidate knowledge and skills related to English language proficiency, English reading fluency, English vocabulary development and guided reading. Flashcards, Big Books and Posters are provided to teachers to use in an integrated manner when teaching vocabulary, phonics and reading. From Grade 2 onwards, ten titles of graded reading books will be provided for each classroom. The accompanying graded reading materials provide ample material for learners to practice decoding and reading at their level of development. It is recognized in South Africa that the opportunity to learn EFAL may be hindered by a lack of suitable materials to assist in the progression from one phase of reading

acquisition to the next, and this is likely to be particularly true in African language schools, the focus of our intervention. Furthermore, teachers are required to provide regular assessment of learners' reading proficiency in order to assign learners to the appropriate graded readers and small reading groups, based on ability.

In addition to the above, coaching is provided as more intensive training to improve teacher capacity. In essence, the role of coaching support is to fuse capacity building and accountability. The assumption is that, just like learning to read, the ability to teach is a skill that needs to be developed over time and might not be accomplished in one day of training. Furthermore, the coach also plays the role of a 'champion' who keeps the teacher accountable to implementing the programme through encouragement and motivation.

3.3.3. Intervention 1 specifically

In intervention 1, specialist reading coaches visit each intervention school about once a month. The in-class support allows for modelling of the new practice on site and the gradual development of teachers in the new practice from novice to expert. The in-class support also allows teachers to manage the emotional labour, i.e. stress, insecurity and anxiety associated with developing a new professional practice mid-career. The presence of the in-class support allows for the development of professional accountability in an environment of trust, where the coach monitors and evaluates the teachers' teaching practices in order to encourage more productive teaching practices. The on-going support from the coach also encourages the teacher to keep up with the increased pace of the lesson plans throughout the course of the year.

3.3.4. Intervention 2 specifically

The technology-supplemented intervention aims to provide new forms of support and guidance on teaching strategies through a range of materials, teaching guides, videos and interactive support platforms that are available at all times to the teacher. These resources are intended to encourage more productive teaching practices among the teachers. Lesson plans are provided in an electronic format on a tablet and are integrated with various audio and visual resources to support teachers in the teaching of EFAL. The resources are supplemented with virtual coaching through phone calls and cell phone messaging. Similar to intervention 1, the on-going support from the virtual coach intends to encourage the teacher to keep up with the increased pace of the lesson plans throughout the course of the year. However, given that the virtual coach will not be in the classrooms, the role of the coach will focus more on encouragement and less on monitoring.

3.4. Evaluation

The study is designed as a randomised control trial and consists of an implementation and evaluation side. Implementation service providers were appointed to ensure the effective implementation of the interventions. The evaluation component of the study is conducted

independently of the implementation service providers to ensure that the impact of the interventions to minimize service provider bias. Data collection service providers will be recruited for each round of data collection and the Project Management Team is ultimately responsible for the instrument design and data analysis.

3.4.1. Sample selection and intervention assignment

Through a process of elimination we developed a sampling frame of 180 eligible schools. Beginning with 731 primary schools registered in the 2016 administrative data in the districts of Ehlanzeni and Gert Sibande, we firstly excluded relatively affluent schools (those in quintiles 4 and 5). Next, we excluded schools in which the language of instruction in the Foundation Phase was neither Siswati, nor isiZulu. We also excluded schools which were missing in the 2014 Annual National Assessment (ANA) dataset.² We further excluded particularly small schools (fewer than 30 Grade 1 enrolments) since many of these schools would practice multi-grade teaching rendering the grade-specific lesson plans less appropriate. We also excluded particularly large schools (more than 160 grade 1 enrolments, or more than three classes in Grade 1, or classes with more than 60 learners in) to limit intervention costs. After all of these exclusions 193 eligible schools remained. Using a random number generator, we then excluded three further schools to remain with a sample of 190 schools. The 190 school sample included a sample of ten replacement schools (one in each of the strata) should the need arise to drop one of the sample schools. We thus obtained the sampling frame of 190 schools.

To increase power and assure balance between intervention arms, we performed stratified randomisation. We created ten strata of 19 similar schools based on school size, socio-economic status, and previous performance in the Annual National Assessments. Within each stratum, we then randomly assigned five schools to each intervention group, eight to the control group and one as a replacement school. Thus we randomly assigned 50 schools to each intervention and 80 to the control. Given that we aim to collect data on 20 Grade 1 learners per school, this sample should be sufficient to identify a minimum effect size of 0.21 standard deviations when comparing an intervention group with the control group and a minimum effect size of 0.23 standard deviations when comparing two intervention groups. These calculations assume a 95% confidence interval, an alpha value of 0.05, an intra-class correlation coefficient (rho) of 0.3 and a correlation between pre- and post-test scores of 0.7.

3.4.2. Instrument Design

The learner assessment was designed in collaboration with a test development expert, Dr Carol Macdonald. The test was designed on the premise that certain basic literacy skills acquired in a first language are transferred to a second language, and therefore the test comprised of both

² The Annual National Assessments have not been administered since 2014. This is therefore the most recent standardised measure of school performance we have for the full population of schools.

home language (in this case isiZulu and siSwati) and English items. Given that the baseline data collection was to be conducted at the start of Grade 1, the test was designed to test learners' home language literacy skills at the start of Grade 1, as well as skills that are predictive of future literacy development. Furthermore, since the interventions are targeting English as First Additional Language as an outcome, the test included a few basic English vocabulary items to determine learners' familiarity with the language at the start of school. The main purpose of the assessment was not to benchmark learner performance against curriculum requirements, but rather to determine learners' literacy abilities at the start of Grade 1. To this end it was important to ensure that the test discriminated well among children at all levels of proficiency.

The test was designed to be orally administered by the fieldworkers, and to be captured electronically on the Tangerine software. In order to test the targeted 20 learners within one school day, the test was designed to take no longer than 15 minutes to administer. The test included various sub-tasks including word recall, non-word recall, phonemic awareness, letter sound recognition, word recognition, listening comprehension, sentence reading and English vocabulary. The test was originally designed in English and then versioned to SiSwati and isiZulu. In the versioning, specific care was taken to use words which are similar in the two languages, to minimise any bias that can be introduced through the language use in the assessment. The instrument was first piloted among Grade R and Grade 1 learners in an isiZulu school in Gauteng, after which the second version was also piloted in five isiZulu and SiSwati schools in Mpumalanga.

The contextual questionnaires were designed by adapting the instruments which were administered in the Early Grade Reading Study in North West to the context in Mpumalanga. An additional focus on the frequency of using English both at home and in the classroom were included in both the Home Background Questionnaire and the Teacher Questionnaire.

3.4.3. Data Collection

The first wave of data collection took place in February 2017 and four subsequent waves of data collection are planned for the duration of the study: wave 2 at the end of Grade 1 (2017), wave 3 at the end of Grade 2 (2018), wave 4 at the end of Grade 3 (2019) and wave 5 at the end of Grade 4 (2020). In each round of data collection, the exact same learners will be identified and assessed to allow for the longitudinal tracking of learners over the duration of the study.

The baseline data collection was conducted by an external service provider which employed fifteen pairs of fieldworkers over a three week period to test of a random sample of 20 learners in each school. The data collection at each school was conducted by a pair of two fieldworkers with one fieldworker being responsible for the individual administration of the learner assessment, and the second fieldworker being responsible for the structured questionnaires. The fieldworker responsible for the learner assessment was also responsible for arranging the completion of a structured contextual questionnaire (Home Background Questionnaire) by the parents or caregivers of all the sampled learners. The contextual questionnaires were taken

home by the learners and fieldworkers arranged to retrieve them from the schools again via the learners' teachers within the three week fieldwork period. The second fieldworker was responsible for administering a structured questionnaire to all the Grade 1 teachers, as well as the school principal. Both these instruments also involved completing some school and classroom facility observations.

The learner assessment, teacher and principal questionnaire were administered using the Tangerine software. The home background questionnaire was sent home with the learners for a parent or guardian to fill out. These questionnaires were subsequently collected by the data collection company and captured and cleaned. The final data set, as well as .do files used to clean the data in the statistical software, Stata, was provided to the project management team at the conclusion of the data collection contract.

During the first week of the baseline data collection the South African Democratic Teacher Union (SADTU) held a protest which disrupted schooling in the Ehlanzeni district. To mitigate the effects of the protest action, the fieldwork schedule was re-arranged to visit affected schools later in the data collection periods. In some schools the protest action took the format of "go-slows" where the schools only operated until 10:00 in the morning. In these schools arrangements were made with the principal and teachers to interview them after 10:00 and to test as many learners as possible before the school closed at 10:00. The protest action has resulted in a slightly smaller sample of learners than expected. Table 1 below shows that the protest action did not affect any of the groups of schools more severely than the rest.

Table 1: Percentage of learners tested during baseline data collection

	Original Number	Tested Number	Percentage Tested
Control	1,600	1,555	94%
Intervention 1	1,000	952	95%
Intervention 2	1,000	975	98%
Total	3,600	3,482	97%

The teacher questionnaires were linked to the learner assessments and the principal questionnaires by means of a linking form that was filled out by the fieldworkers. Unfortunately, the linking forms were filled out very poorly, specifically with regards to the teacher names and surnames. This has meant that the linking of the teacher questionnaire to the learner questionnaires has been challenging, and only 89% of the teacher interviews could be linked to the learner assessments. The linking of the teacher questionnaires to the learner assessments were also more effective among the intervention 1 schools, and therefore introduces a potential bias in the analysis of the results when including the teacher interview data. Robustness checks will be conducted to evaluate the impact of the bias.

Table 2: Percentage of teacher interviews conducted during baseline data collection

	Number of teacher interviews	Number of teacher exercises	Number of learners matched to teachers	Percentage Questionnaires merged
Control	143	141	1225	84%
Intervention 1	89	86	895	97%
Intervention 2	88	88	843	89%
Total	320	315	285	89%

The principal interview asks general questions about the school to gain a better understanding of the context in which the Grade 1 teachers and learners function. Fieldworkers were briefed that the deputy principal or a Head of Department (HOD) in the school can also respond to the questionnaire, should the principal not be available on the day of the visit. Table 3 shows that 78% of the interviews were conducted with the principal themselves, but that 40 of the interviews were conducted with a School Management Team member (SMT) other than the principal.

Table 3: Number of principals interviewed during baseline data collection

Respondents	Number
Principal him/herself	140
On behalf of the Principal	40

The final instrument administered was the home background questionnaire which learners took home. The purpose of the questionnaire was to collect information about the home circumstances the learner comes from. Since the baseline data collection was administered at the start of Grade 1, it is expected that home background factors will play a large role in determining learner performance. The data collection company collected 3,482 returned questionnaires, of which 3,327 (96%) were successfully matched.

Table 4: Percentage of Home Background Questionnaires returned

	Number returned	Number matched	% Learner matched to PBQ
Control	1,555	1,459	94%
Intervention 1	952	924	97%
Intervention 2	975	944	97%
Total	3,482	3,327	96%

3.5. Ethical Clearance

The research methodology, with the intended instruments and research consent forms related to the baseline data collection, was formally submitted to the University of Witwatersrand Human Research Ethics Committee (Non-medical) in May 2016. The project was approved on 3 November 2016 and was granted ethical clearance for the duration of the study, up to 2 November 2019. During the baseline data collection, consent forms were provided to all principals and teachers interviewed. Consent forms were also sent with the home background questionnaires to parents.

4. Baseline Results

The sections to follow report on the data collected through the baseline data collection. The first section describes the learner characteristics of the learners in the sample, as well as the teacher and school characteristics linked to the learners. The second section reports on the baseline learner assessment and the learner performance on the various tasks at the start of Grade 1. Subsequent to the learner assessment analysis, balance tests are reported to ensure that the sample selected and the intervention allocation support the assumption that the random allocation of the schools to the intervention groups has ensured that there are no observable differences to the groups of schools at the baseline. Finally some sensitivity checks are conducted to evaluate the effect of various data collection and intervention implementation issues so that they do not cause any reason for biased results.

4.1. Sample Characteristics

As mentioned above, the sample was specifically selected to target the lowest performing portion of the population. From Table 5 it is evident that over half of the schools in the sample are Quintile 1³ schools, whereas almost a third of the schools are Quintile 2 schools. Only 13% of the schools in the sample are Quintile 3 schools. There are slightly more schools in the Ehlanzeni district relative to the Gert Sibande district, and there are remarkably more schools with SiSwati as the language of learning and teaching, than isiZulu schools. The administrative data which classifies school as either being located in urban or rural areas are of a poor quality for the Mpumalanga province, and therefore school principals were asked to select the area in which their schools are located. 74% of school principals reported that their schools were situated in a remote rural area, a further 18% stated that their schools are in a small village and the final 8% of principals reported that their schools were either in an informal settlement, a formal township or a suburban area.

³ The Quintile system in South Africa determines the funding that schools are eligible for. Quintile 1 schools are situated in the poorest areas, whereas Quintile 5 schools are situated in the wealthiest neighbourhoods. Quintile 1 – 3 schools are classified as no-fee schools, and are therefore not allowed to charge learners any school fees.

Table 5: School Characteristics

	Number of schools	Percentage of schools	Percentage of learners
Quintile 1	97	53.9%	54.2%
Quintile 2	59	32.8%	32.6%
Quintile 3	24	13.3%	13.2%
Ehlanzeni	98	54.4%	52.9%
Gert Sibande	82	45.6%	47.1%
isiZulu	49	27.2%	28.6%
SiSwati	131	72.8%	71.4%
Remote rural	134	74%	73.7%
Small village	32	18%	18.2%
Informal settlement	3	2%	1.7%
Formal Township	9	5%	5.3%
Suburban	2	1%	1.1%

4.1.1. Learner Characteristics

The largest majority of learners in the selected sample are six years of age, which is the expected age at the start of the Grade 1 school year. The South African Schools Act specifies that a learner can be enrolled in Grade 1 if the learner is five years and turning six years or older by 30 June of the Grade 1 year. At the start of the Grade 1 year it is evident that only 17% of the learners are actually five years or younger, whereas a quarter of the learners are already seven years or older. The learners who are seven years or older may have been enrolled a year later, but it is likely that they are learners who are repeating Grade 1. The sample includes slightly more boys than girls, but the difference is not relevant. Two-thirds of the learners' mothers filled out the home background questionnaire as their main caregiver and 18% of the learners' grandmothers filled out the questionnaire as a caregiver. Smaller proportions were filled out by fathers, siblings or other household members. The questionnaire specifically asked the age of the learner's mother, but it is likely that this question might have been misinterpreted by the person filling out the questionnaire in that they filled out their own age. To mitigate some of this misinterpretation, mother age was calculated only for those individuals who reported that they are the learner's mother, as well as all other individuals reporting an age which is above 15 and below 45. Using this variable, it is evident that 12% of learners are from households where the mother is younger than 25 (which would mean that the learner's mother would have been younger than 20 when they were pregnant with the learner). With regard to parent/ guardian education, 56% of parents/ guardians did not complete matric, whereas a further 23% only completed matric. This means that 79% of the learners had a parent/ guardian who did not have any post-school education. A further 7% of learners are from households where the parent is studying towards, or have completed a certificate or diploma, and only 6% are studying towards, or have completed a degree.

Table 6: Learner characteristics

		# learners	% learners	Control	Intervention 1	Intervention 2
Learner Age	<i>Younger than 5</i>	46	1%	1%	1%	1%
	<i>5 Years Old</i>	546	16%	17%	14%	17%
	<i>6 Years Old</i>	1,912	57%	58%	61%	53%
	<i>7 Years Old</i>	697	21%	20%	19%	23%
	<i>Older than 7</i>	126	4%	3%	4%	4%
Learner Gender	<i>Boys</i>	1,801	54%	53%	54%	55%
	<i>Girls</i>	1,526	46%	47%	46%	45%
Main Caregiver	<i>Mother</i>	2,190	66%	64%	66%	68%
	<i>Father</i>	203	6%	6%	6%	6%
	<i>Grandmother</i>	583	18%	18%	19%	16%
	<i>Sibling</i>	120	4%	4%	3%	3%
	<i>Other</i>	231	7%	7%	7%	6%
Mother age	<i>Younger than 25</i>	402	12%	12%	13%	12%
	<i>25 - 34 Years Old</i>	1,638	49%	49%	50%	48%
	<i>35 - 44 Years Old</i>	722	22%	21%	22%	22%
	<i>Older than 45</i>	134	4%	4%	3%	5%
	<i>Missing</i>	431	13%	13%	12%	13%
Mother Education	<i>Did not complete matric</i>	1,875	56%	55%	58%	56%
	<i>Finished Grade 12</i>	768	23%	22%	24%	24%
	<i>Studying towards TVET/ FET certificate</i>	102	3%	4%	2%	3%
	<i>Finished a certificate/ diploma</i>	118	4%	4%	3%	3%
	<i>Studying/ completed a degree</i>	184	6%	5%	6%	5%

Parents or guardians of the learners were also asked some questions on how regularly they read to their child, or speak English to their child. They were also asked some questions on household assets so that a wealth index could be derived for each learner. Self-reported answers on language practices in the home should be interpreted with some caution, as these types of questions are susceptible to respondents providing socially desirable responses. 60% of the parents or guardian in this sample reported that they sometimes speak English to their children, but 27.5% reported that they never speak English to their child. With regards to reading to their children, 35.9% of parents responded that they read to their child every day, whereas 27.9% responded that they read most days and 22.9% that they read once a week. Parents were also asked who they think is most responsible for their child's reading progress. Almost two-thirds of the parents said that they regard themselves as most responsible for their children's reading progress, 27% responded that teachers are most responsible and a small percentage said that it is the government's responsibility.

Figure 2: Who is responsible for your child's reading?

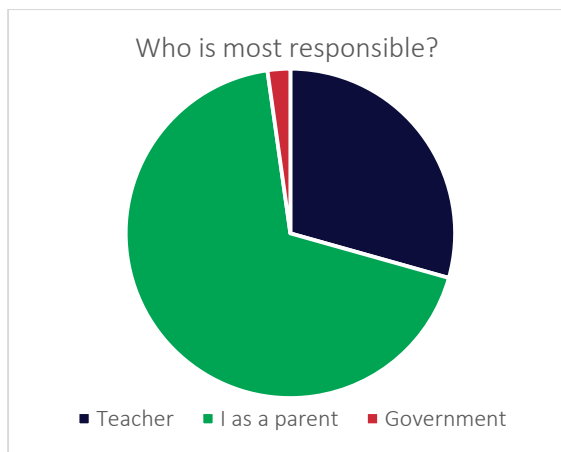


Figure 3: How often do you speak English to your child?

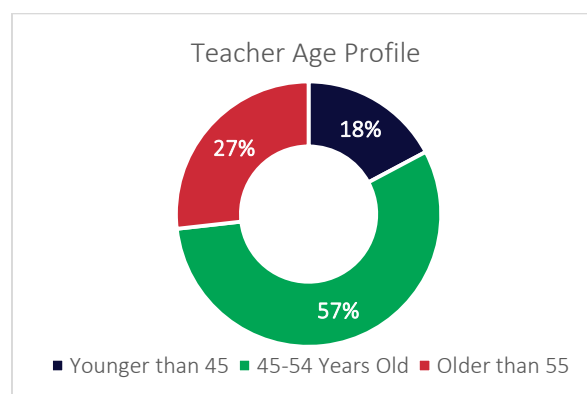


4.1.2. Teacher Characteristics

Given the poor linking between the learners and the teachers in the sample, the teacher characteristics were calculated on the teacher dataset separately. For this reason the statistics reported on in this section should be interpreted as a proportion to the teachers interviewed, and not to the proportion of learners in the sample. In total, data was collected through interviews with 320 teachers.

The teacher age profile reveals that more than a quarter of the Grade 1 teachers are older than 55 years, which means that one in every four Grade 1 teachers will be retiring in the next ten years. Teacher education levels indicate that the majority of the teachers have a three or four year diploma and that only 18% of the teachers have only a three-year diploma.⁴

Figure 4: Teacher age profile



Some questions were also asked regarding the teachers' current planning and teaching practices with regards to the teaching of English as First Additional Language. The overwhelming majority of teachers reported that they use the EFAL CAPS document and the DBE workbooks to plan their EFAL teaching. Very few teachers reported using plans they developed themselves, or even plans developed by another teacher. Teachers were further

⁴ Teacher education levels are not shown in a table or graph as the categories are were not asked as being mutually exclusive. Regardless, it is possible to deduct that the majority of teachers have obtained at least a three or four year diploma.

asked which lesson types they consider most important in the teaching of EFAL in Grade 1 (the CAPS curriculum specifically states that Listening and Speaking, and Shared Reading are the most important). 77% of teachers correctly identified Listening and Speaking as one of the most important lesson types, but only 48% managed to correctly identify Shared Reading. Teachers were also asked which of the following three subjects they enjoyed teaching most in Grade 1: Mathematics, EFAL or Home Language. Interestingly, the majority of teachers said that they enjoyed teaching Mathematics most, followed by Home Language. EFAL was the subject that teachers were least likely to state as the subject they enjoyed teaching most.

Figure 5: Resources used for EFAL lesson planning

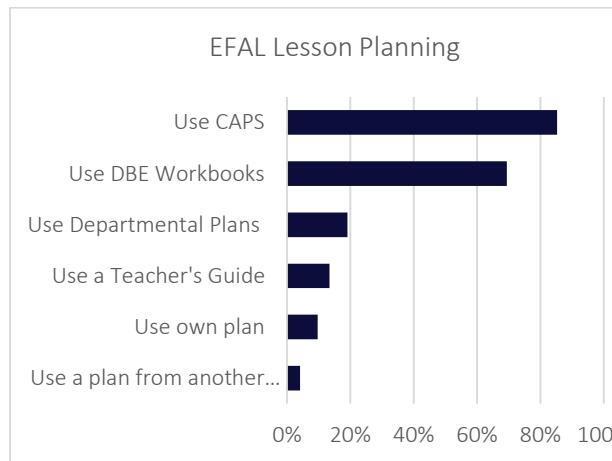


Figure 6: Two most important lesson types

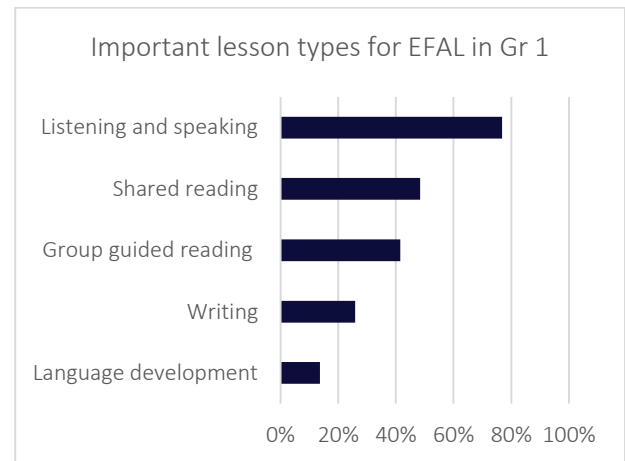
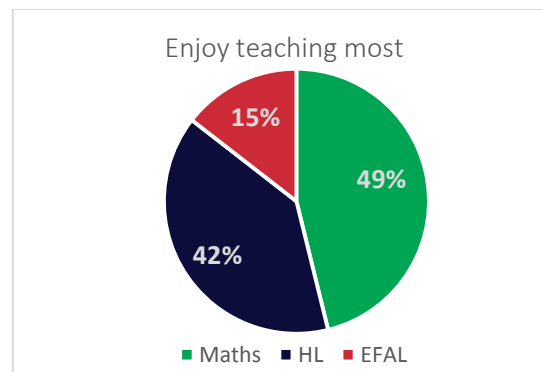
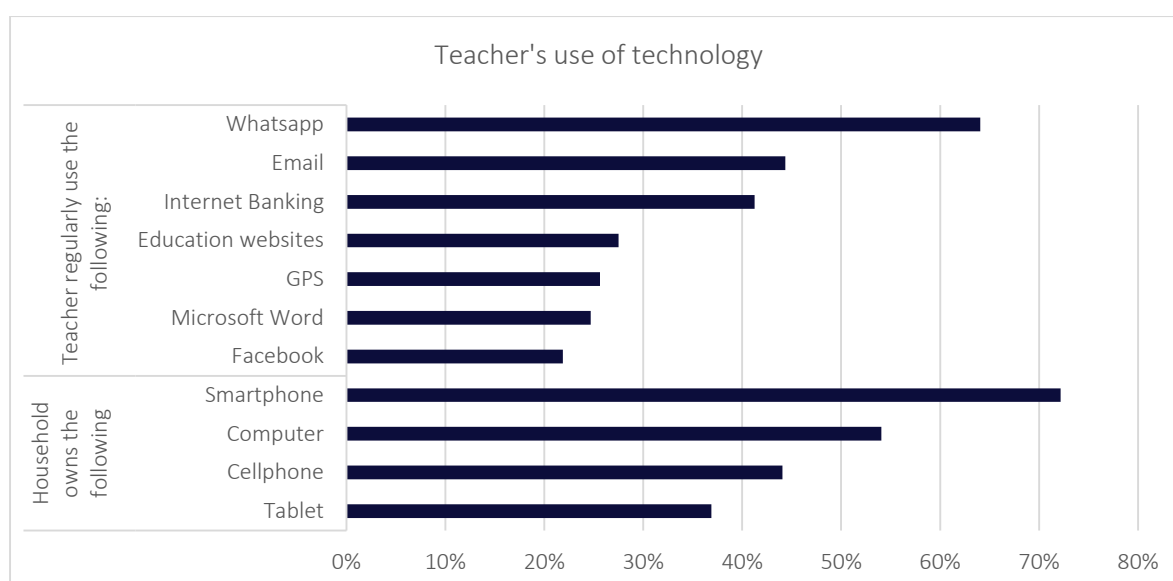


Figure 7: Subject teachers enjoy teaching most



Given the ICT focus of intervention 2, all teachers were asked about the technology devices they have in their household, as well as how regularly they use certain technology features. It was interesting to note that 72% of teachers owned a smartphone and 54% owned a computer. However, only 44% of teachers made use of email, only 28% accessed educational resources on the internet and only 25% made use of Microsoft Word. However, 64% of teachers use Whatsapp, which is encouraging for intervention 2, as this will be one of the methods of communicating with the virtual coach.

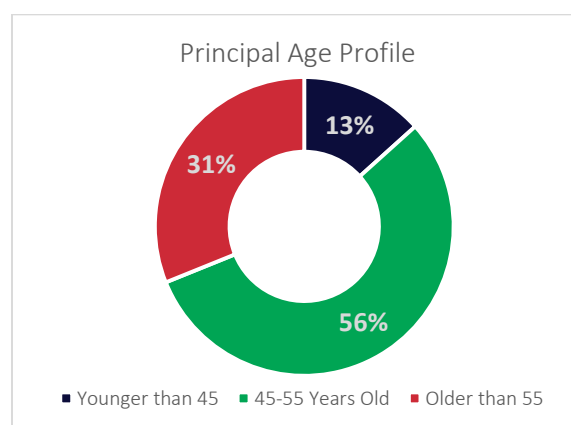
Figure 8: Teacher's use of technology



4.1.3. School and Principal Characteristics

Similar to the teacher age profile, it is evident that a large percentage of principals will be retiring in the near future. 31% of Principals are older than 55 years, which means that almost one in every three principals will be retiring within the next ten years. There is also not a strong gender bias with regards to the principals in the sampled schools, with 46% of the principals being female.

Figure 9: Principal age profile



Two curriculum related questions were asked of the principals in order to gain a sense of the instructional leadership present at the school. The first question asked the principal how many hours a week are dedicated to teacher EFAL in the foundation phase. The CAPS curriculum provides two options, either the minimum amount of two hours, or the maximum amount of three hours. Four options were presented to principals (one to four hours) and interestingly 37% of the principals went for the seemingly most socially desirable option, four hours. The second question asked principals when learners should be starting with EFAL in the foundation

phase. This question was answered more accurately, with 71% of the principals providing the correct answer of ‘at the start of Grade 1’.

Figure 10: Instructional Leadership - Time is dedicated to EFAL?

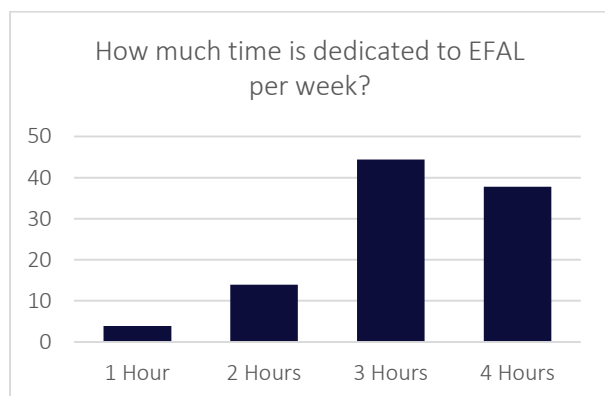
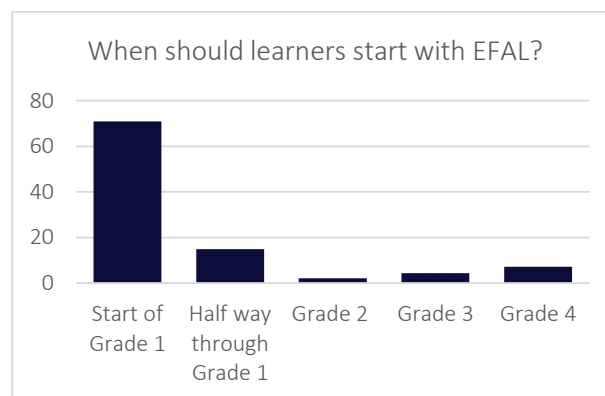


Figure 11: Instructional Leadership -When to start with EFAL?



4.2. Learner Assessment

The baseline learner assessment was adapted from the isiZulu and SiSwati Early Grade Reading Assessment and included measures of pre-literacy skills. The test was therefore administered orally to learners, one learner at a time. The test took on average 15 minutes to administer and included nine different sub-tasks. The standard EGRA items which remained in the test included letter sound recognition, familiar word reading and sentence reading. Given that the assessment was conducted at the start of Grade 1, before learners begin formal literacy instruction, various sub-tasks were included to assess pre-literacy abilities. These include receptive and expressive vocabulary, non-word recall, phoneme isolation, listening comprehension and visual perception. Finally, a task was included to assess the learners' English vocabulary.

The average performance and summary statistics for each sub-task are presented in Tables 7 and 8. Learners scored relatively high in the pre-literacy tasks, but predictably struggled with the Grade 1 EGRA tasks since the majority of learners would have only been exposed to formal literacy instruction for about two weeks. Overall, the test did not seem to have any floor or ceiling effects, but within some of the sub-tasks these effects were seen. Learners performed scored particularly high on the pre-literacy tasks assessing expressive vocabulary and word and non-word recall, with learners at the 10th percentile managing to answer a few items correctly. These items were included purposefully to allow for discrimination in learner performance among the weaker part of the learner performance distribution. The EGRA task assessing word recognition and sentence reading was clearly too difficult for the learners, with quite severe floor effects visible in the performance results on these tasks. This result is expected since learners are at the start of formal literacy instruction. Nonetheless, these items will assist in discriminating amongst the strongest learners. The letter sound recognition and story comprehension items both have a more normal distribution. While learners are not able to read words, they may have been exposed to letters in Grade R or at home. Overall the test

seems to provide enough information to discriminate among learners at various parts of the performance distribution. Using Principal Component Analysis, the performance results on all the different sub-tasks are combined to give a single composite score. Figure 12 shows the learner performance distribution for all learners, and it is encouraging to notice that this distribution has a good normal-curve shape which will allow the detection of a change in learning outcomes for all learners across the distribution. The sections to follow will describe the results of each sub-task in more detail.

Table 7: Average performance per sub-task

	Observations	Mean	S.D.	s.e.
<i>Naming the Animals (/12)</i>	3,327	7.30	2.37	0.04
<i>Word Recall (/14)</i>	3,327	10.00	2.33	0.04
<i>Non-word Recall (/5)</i>	3,327	4.21	1.13	0.02
<i>Phoneme Isolation (/6)</i>	3,327	1.11	1.79	0.03
<i>Story Comprehension (/4)</i>	3,327	2.20	1.16	0.02
<i>Letter Sound Recognition</i>	3,327	4.62	5.01	0.09
<i>Word Recognition</i>	3,327	0.41	1.37	0.02
<i>Sentence Reading</i>	3,327	0.04	0.38	0.01
<i>Visual Perception (/10)</i>	3,327	1.55	1.64	0.03
<i>English Vocabulary (/6)</i>	3,327	0.82	1.03	0.02
<i>Index Score</i>	3,327	0.00	1.00	0.02

Table 8: Summary Statistics per sub-task

	Observations	p10	p25	p50	p75	p90	Min.	Max.
<i>Naming the Animals</i>	3327	4	6	7	9	10	0	12
<i>Word Recall</i>	3327	7	9	10	12	13	0	14
<i>Non-word Recall</i>	3327	3	4	5	5	5	0	5
<i>Phoneme Isolation</i>	3327	0	0	0	2	4	0	6
<i>Story Comprehension</i>	3327	1	1	2	3	4	0	4
<i>Letter Sound Recognition</i>	3327	0	1	3	6	12	0	20
<i>Word Recognition</i>	3327	0	0	0	0	1	0	6
<i>Sentence Reading</i>	3327	0	0	0	0	0	0	6
<i>Visual Perception</i>	3327	0	0	1	3	4	0	10
<i>English Vocabulary</i>	3327	0	0	0	1	2	0	6

Figure 12: Overall learner performance distribution

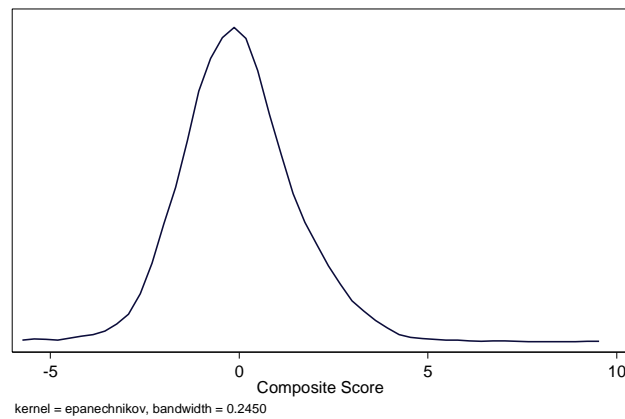


Table 9 presents the correlation matrix for the sub-tasks in the baseline learner assessment. The floor effects in the word recognition and sentence reading assessment means that not enough information on learner performance has been captured in these tasks to calculate credible correlations. As expected, however, correlations were higher between word recall and non-word recall, and letter sound recognition and story comprehension. English vocabulary was more correlated with the first sub-task (naming animals in the home language) and story comprehension, but very there was very little correlation between English vocabulary and the word recall and non-word recall tasks. However, overall the correlations between the sub-tasks are not specifically high and is indicating that there is quite a bit of noise in the baseline data.

Table 9: Sub-test Correlation Matrix

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9	Task 10
Naming Animals	1.00									
Word Recall	0.20	1.00								
Non-word Recall	0.13	0.33	1.00							
Phoneme Isolation	0.13	0.17	0.15	1.00						
Story Comprehension	0.23	0.21	0.16	0.22	1.00					
Letter Sound Recog	0.10	0.10	0.08	0.37	0.13	1.00				
Word Recognition	0.01	0.06	0.01	0.08	-0.02	0.36	1.00			
Sentence Reading	0.05	0.02	0.01	0.11	0.06	0.17	0.17	1.00		
Visual Perception	0.14	0.12	0.12	0.16	0.19	0.09	-0.08	0.03	1.00	
English Vocabulary	0.20	0.09	0.08	0.19	0.21	0.18	0.01	0.08	0.13	1.00

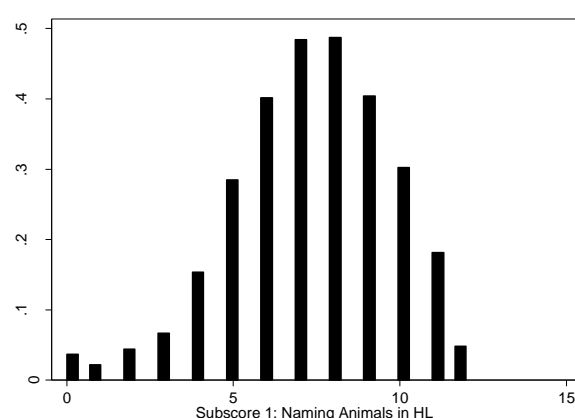
Sub-task 1: Naming Animals

The first sub-task of the learner assessment was an easy task that entailed learners looking at a picture of farm animals and naming each of the animals that they recognised in their Home Language. Overall, learners did quite well in this exercise, with learners in the 10th percentile at least managing to recognise four of the animals. On average, learners managed to name at least seven animals. Learners did, however, struggle to identify the lamb, which might be due to the colouring of the lamb in the picture resulting in the lamb being unclear.

Table 10: Subtask 1 - Naming Animals

	% Correct	s.e.
<i>Cow</i>	93%	0.004
<i>Chicken</i>	93%	0.005
<i>Dog</i>	88%	0.006
<i>Horse</i>	74%	0.008
<i>Duck/Bird</i>	70%	0.008
<i>Pig</i>	63%	0.008
<i>Goat</i>	60%	0.008
<i>Rabbit</i>	55%	0.009
<i>Donkey</i>	39%	0.008
<i>Sheep</i>	38%	0.008
<i>Lamb</i>	6%	0.004
Cronbach:	0.6758	

Figure 13: Histogram – Naming Animals



Sub-task 2: Word and Non-word Recall

Two sub-tasks, word recall and non-word recall were included to assess the learners' phonological working memory in their Home Language. The ability to repeat real and non-words has been closely linked to vocabulary acquisition and growth (Gathercole, 2006). The first task required learners to repeat a sequence of home language words after the fieldworker, with the length of the sequence increasing with each item. The second task required learners to repeat non-words which follows the same language structure as the home language, with the words increasing in complexity in each item. Learners did quite well in this exercise, with the majority of learners managing to repeat on average ten of the fourteen words, and four of the five non-words. Learners at the 10th percentile managed to recall at least seven of the fourteen words, and three of the five non-words.

Table 11: Sub-task 2.1 Word Recall

isiZulu/ SiSwati	% Correct
<i>Cula / Cula</i>	95%
<i>Jika/Jika</i>	97%
<i>Vela/Vela</i>	95%
<i>Xoxa/Coca</i>	88%
<i>Qala/Cala</i>	88%
<i>Sika/Sala</i>	91%
<i>Itiye/Umlimi</i>	79%
<i>Indlebe/Indlebe</i>	79%
<i>Baleka/Baleka</i>	85%
<i>Indoda/Indvodza</i>	58%
<i>Izitsha/Inkhomo</i>	46%
<i>Umhlobo/Umhlobo</i>	22%
<i>Incwadi/Incwadzi</i>	35%
<i>Umntwana/Umntfwana</i>	45%
Cronbach:	0.7496

Figure 14: Histogram - Word recall

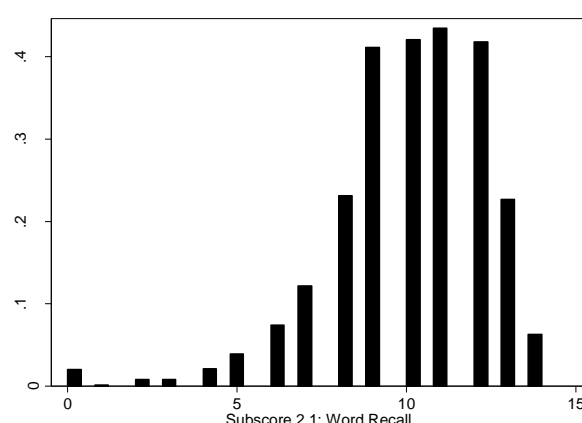


Table 12: Sub-task 2.2. Non-word Recall

isiZulu/ Siswati	Observations	% Correct	s.e.
Mani/Mani	3,327	96%	0.004
Inxhovi/Inchoti	3,327	89%	0.005
Zonhlamvu/Temdzalu	3,327	85%	0.006
Noyiqoqo/Ofudlanga	3,327	77%	0.007
Qinisokisa/Ngiyagudziye	3,327	74%	0.008
Cronbach:	0.626		

Sub-task 3: Phoneme Isolation

Phonological awareness is widely accepted as one of the most important cognitive skills for the acquisition of literacy. Phonological awareness is the ability to identify and manipulate the phonological structures (sounds) in words independent of their meaning. This skill is a pre-requisite for reading as it enables learners to match letters to the sounds they represent. Learners who are slower at matching sounds to letters will understandably take longer to learn to read fluently. Although a large body of research exists for assessing the phonological awareness of learners in English, very little research exists on the assessment of this skill in African languages in general, but specifically the Nguni language group to which both isiZulu and SiSwati belong. Diemer, et al. (2015) recently conducted research on the development of phonological awareness measures for isiXhosa and advocate the need for further research into phonological awareness measures in African languages. In the baseline learner assessment, phoneme isolation was used to evaluate the learners' phonological awareness. This sub-task required learners to provide the first sound of a word which was read out to the learner by the fieldworker. The ability to identify phonemes may emerge before the start of formal reading instruction, but generally it is accepted that this skill only develops further after learners are exposed to reading instruction (Adams, 1990). The latter may be the case in South Africa where there is not much research on this skill, and therefore less information for parents on how they can develop their child's (pre) literacy skills in the home language. The words used in this task were words which are similar in SiSwati and isiZulu, and therefore it is not expected that there will be any differences in learner achievement between the two languages.

When the fieldworkers administered this task, learners often only repeated the whole word after the fieldworker. It is, unfortunately not clear whether this repetition was due to the learners not understanding the task correctly and were therefore still using the same 'rules of the game' from the previous exercise (word recall/ non-word recall), or whether it was that the learners just could not identify the initial sounds in the words.

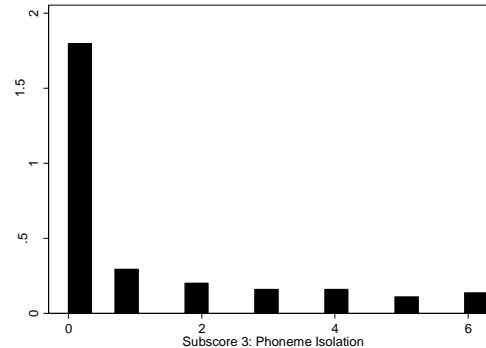
Regardless of potential errors from the administration of the task, learners performed as would be expected for first graders: 63% of the learners could not identify a single sound, but some learners (about 13%) could identify three or more sounds. The average score out of six was

1.11. From Table 8 it is evident that it is only at the 75th percentile that learners managed to successfully identify a phoneme. The manner of articulation of the sounds to be identified did not affect the learners' ability to identify them.⁵

Table 13: Sub-task 3 - Phoneme Isolation

	% Correct	s.e.
/b/uka	20%	0.007
/g/ogo	22%	0.007
/w/ena	13%	0.006
/c/eba	22%	0.007
/dl/ala	17%	0.007
/ch/aza	16%	0.006
Cronbach:	0.864	

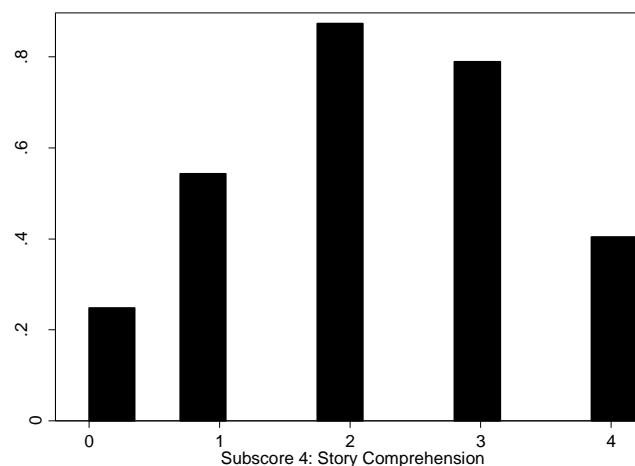
Figure 15: Histogram - Phoneme Isolation



Sub-task 4: Story Comprehension

Sub-task 4 was a story comprehension that entailed the fieldworkers reading a short (five sentence) story to the learners in their Home Language, and asking four factual questions on what happened in the story. Fieldworkers were trained to only read the story to the learners once, and to read the story with expression. In general the learners did relatively well in this sub-task with 81% of the learners successfully answering the first question of the task correctly. On average learners managed to answer two of the four questions correctly, and 14% of learners managed to answer all of the questions correctly.

Figure 16: Histogram - Story Comprehension



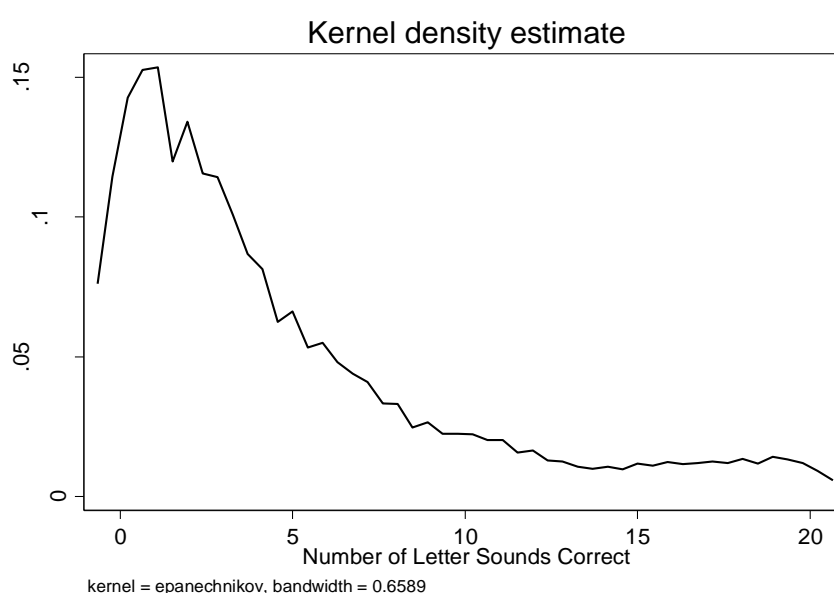
⁵ Typically for English, the easiest sounds to identify would be fricatives ("dl"), sonorants ("w"), followed by plosives ("b"). The jury is still out as to how this will apply to African languages.

Sub-task 5: Letter Sound Recognition

Letter-sound recognition is a key predictor of later reading success especially for first grade English learners (NICHHD, 2000). This task can also be used to quantify learners' previous print and reading instruction exposure: higher print and reading instruction exposure would result in more letters recognised.

The letter sound recognition assessment is based on the letter sound recognition assessment in the Early Grade Reading Assessment (EGRA). Given that the assessment was conducted at the start of Grade 1, it was adapted to include fewer letters, the letters which learners are most likely to be familiar with at the start of Grade 1, and to only be administered for 40 seconds (instead of 60 seconds). On average learners identified four to five letters in the 40 seconds, which translates to about six to seven letters per minute. Learners at the 10th percentile did not manage to identify a single letter, whereas learners at the 25th percentile managed to identify one letter and learners at the 90th percentile, twelve letters.

Figure 17: Distribution of learner performance on Letter Sound Recognition



Sub-task 6 and 7: Word Recognition and Sentence Reading

Sub-tasks 6 and 7 assess learners' ability to read words, in isolation and in sentences. Although it is not expected that learners will be able to decode or recognise words at this stage of their formal education, the task was included to identify learners at the upper distribution of performance, especially since previous reading skill accounts for unique variance in later reading skill.

The learner assessment included six words which learners were asked to read. The words were no longer than four phonemes, and all followed the cvcv⁶ sound pattern which is characteristic

⁶ Consonant – vowel – consonant – vowel, for instance yena or bona.

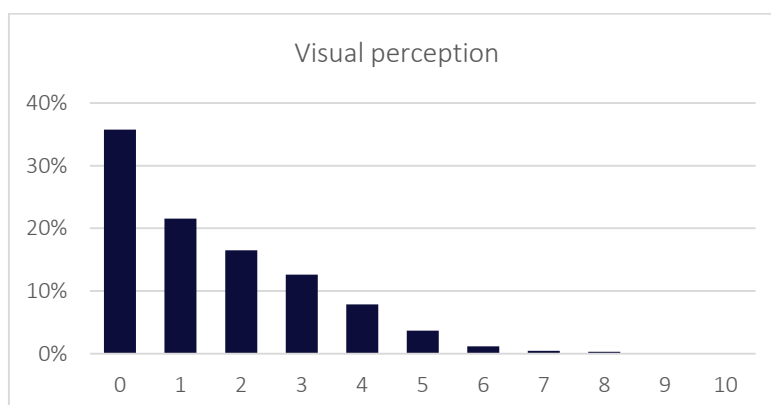
of the Nguni language group. Once again, the words in the different language versions of the task were carefully chosen to be words which are very similar across the languages. Overall, the learners did very poorly in this task, with only learners in the 90th percentile managing to successfully read one word.

The sentence reading task required learners to read two short sentences which were assumed to be very familiar to learners at the start of Grade 1. Similar to the word recognition, learners scored very poorly in this sub-task, with even learners in the 90th percentile not managing to read a single word in these sentences.

Sub-task 8: Visual Perception

The penultimate task of the learner assessment comprised of a 'spot-the-difference' task where learners were requested to identify the differences between two seemingly similar pictures. The task was included as an easier, fun, exercise, with the intention to boost the learners' confidence a bit before the final English task. However, the learners struggled a lot more with this exercise than expected. Over a third of the learners did not manage to identify a single difference. Learners at the median managed to identify one difference, and learners at the 90th percentile managed to identify four differences.

Figure 18: Sub-task 8 - Visual perception



Sub-task 9: English Receptive and Expressive Vocabulary

The purpose of the final task was to assess learners' familiarity with English at the start of Grade 1. The learners were given the same picture of farm animals as the picture used in the first task of the learner assessment. To assess the learners' receptive vocabulary, the first three questions required the fieldworkers to ask learners to point to specific animals in the picture. The instructions for the task was provided in the learners' home language, but the animal name was read in English. For example: '*Ngikhombe dog*'.

The data shows that learners are not yet familiar with English animal names, and it can be assumed, English in general. Just less than a third of learners managed to correctly point to the first two animals (dog and chicken), but only 10% could point to the goat. The task assessing expressive vocabulary entailed the fieldworkers pointing to specific animals and asking the

learners to name the animal in English. Very few learners could correctly name the animals in English, with only 5% of learners managing to name the cow, 3% the pig and 1% the sheep. These results suggest that at the start of Grade 1, learners are not yet familiar with basic farm animals in English.⁷ This in turn gives an indication of learners' general familiarity of English at the start of school.

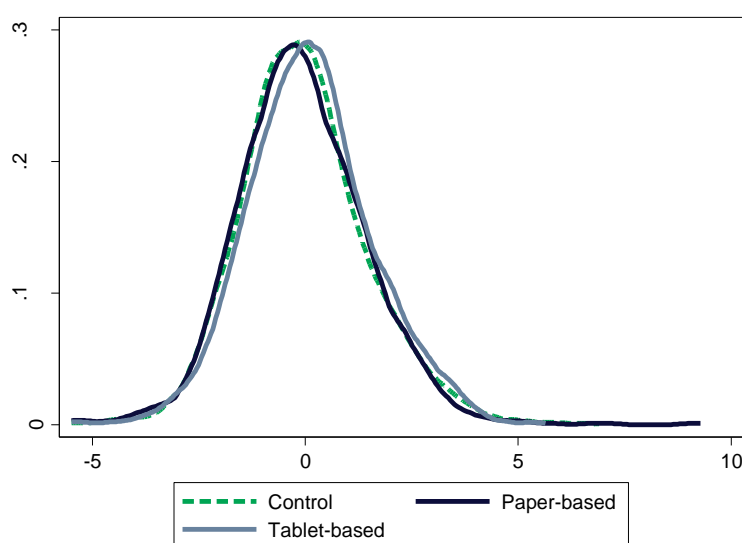
Table 14: English Receptive and Expressive Vocabulary

		Observations	% Correct	s.e.
Receptive Vocabulary	<i>Dog</i>	3324	32%	0.008
	<i>Chicken</i>	3324	30%	0.008
	<i>Goat</i>	3324	11%	0.005
Expressive Vocabulary	<i>Cow</i>	3324	5%	0.004
	<i>Pig</i>	3324	3%	0.003
	<i>Sheep</i>	3324	1%	0.002

4.3. Balance Test

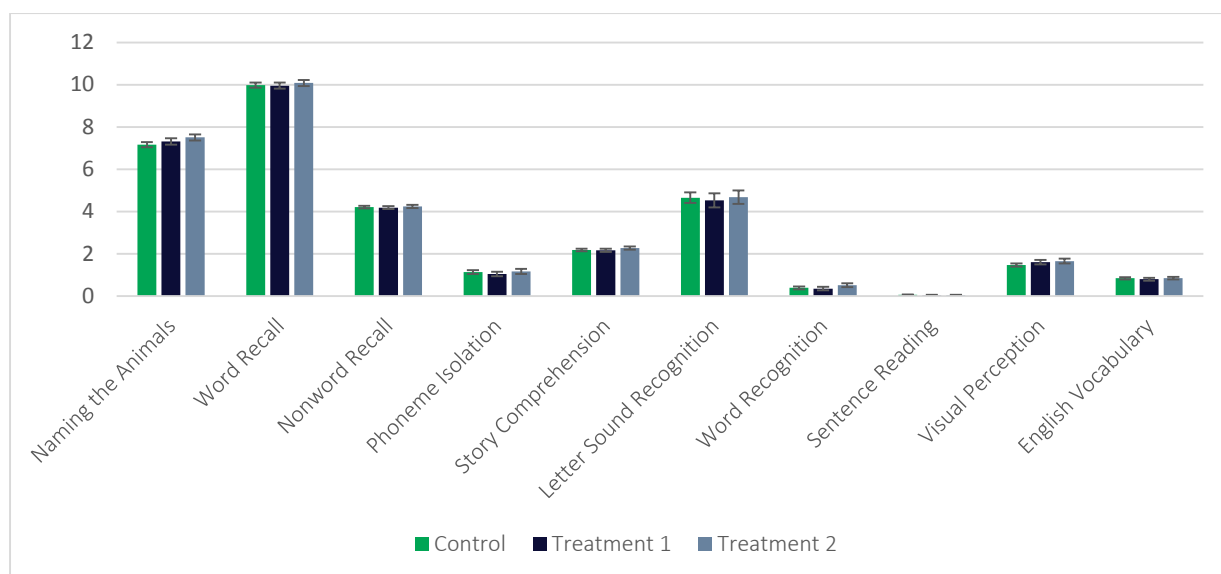
The main purpose of randomisation is to ensure that neither the intervention groups nor the control groups are inherently more prone to showing learning gains over the period of the study. Since this is the fundamental assumption upon which the randomised control trial is based, it is necessary to check that the different groups of schools are balanced. As a first check, Figure 19 shows the learner performance distributions of the different intervention groups. From this figure it is evident that there are no distinguishable differences between the distributions. As a second check, Figure 20 shows the average learner performance in the different sub-tasks for each intervention group, also with no significant differences.

Figure 19: Composite learner score by intervention group



⁷ Farm animals are introduced in the second term of Grade 1 in the DBE workbook.

Figure 20: Sub-task performance by intervention group



Furthermore, balance tests were conducted on each sub-task to ensure that the sample is balanced. Table 14 shows the results based on regression analysis to test for balance. These tests evaluate whether the differences in learning outcomes among the different intervention groups are statistically significantly different from zero. Each column in Table 14 is a separate regression run for each sub-task on intervention indicators, controlling for strata and district fixed effects. The significant differences in the table will be represented by stars. For instance, the one statistically significant result in Table 14 is the performance on the 'naming the animals' sub-task of the learners in the intervention two group, relative to the control group. The two stars indicate that this difference is significant at a 5% level. The final row in the table shows the p-value for the pair-wise test comparing the means between the two intervention groups (i.e. not the means of the intervention groups with the control group). A p-value of less than 0.05 would indicate that there is an imbalance between the two groups for the specific learning outcome. There are only two slight imbalances visible, but overall the samples are clearly balanced.

Table 15: Balance tests per sub-task

	Animals	Word Recall	Non-word Recall	Phoneme Isolation	Comprehension	Letter Sounds	Words Correct	Sentence Reading	Visual Perception	English Vocabulary
Paper-Based	0.15 (0.152)	-0.02 (0.118)	-0.02 (0.067)	-0.06 (0.12)	-0.02 (0.064)	-0.16 (0.484)	-0.05 (0.136)	-0.03* (0.016)	0.17 (0.121)	-0.04 (0.066)
Tablet-Based	0.39*** (0.149)	0.12 (0.118)	0.03 (0.094)	0.04 (0.131)	0.1 (0.063)	0.07 (0.48)	0.14 (0.165)	-0.02 (0.016)	0.19 (0.124)	0.01 (0.057)
Control Mean	7.155	9.981	4.208	1.129	2.179	4.652	0.387	0.051	1.46	0.836
N	3,327	3,327	3,327	3,327	3,327	3,327	3,327	3,327	3,327	3,327
Paper=Tablet: p-value	0.158	0.245	0.573	0.448	0.092	0.66	0.27	0.518	0.857	0.511

Note: Each column represents a separate regression on intervention dummies and stratification dummies. Standard errors are clustered at school level. * for $p < .1$; ** for $p < .05$; *** for $p < .01$

A short English proficiency task was administered to teachers. The main purpose of the task is to serve as a control in the learner regressions, but it also allows an additional balance check. Table 15 shows that the teachers in the various intervention groups are similar in their English proficiency. Although teacher English proficiency is not the focus of the interventions, it is likely that improved English proficiency might be the result of increased use of English during the lessons. The balance between the groups at the baseline means that any differences in teacher English proficiency in the wave 2 data collection, could be ascribed as a secondary outcome of the interventions.

Table 16: Balance test on Teacher English Proficiency

	Teacher Score
Paper-Based	-0.4 (0.256)
Tablet-Based	0.14 (0.267)
N	2963
Paper=Tablet: p-value	0.063
<i>Note:</i> Teacher score regressed on intervention and stratification dummies. Standard errors are clustered at school level. * for $p < .1$; ** for $p < .05$; *** for $p < .01$	

4.4. Which characteristics are correlated with learner test scores?

Table 16 shows the results of simple Ordinary Least Squares (OLS) regressions predicting the overall composite test score, as well as the number of letters read correctly. The overall composite score is reported in terms of standard deviations, whereas the letter sounds are reported in letters read correctly within the 40 seconds time span. The regressions indicate that age is a significant predictor of learner performance, with learners who were five years old or younger at the start of Grade 1 performed significantly poorer than their six year old peers. Learners who were seven years old, in turn performed better than their six year old peers. The language in which the learner took the test seems to have made a potential difference in the letter sound recognition. This is an interesting finding, given that the letter sound recognition task did not differ between the two languages and can therefore be regarded as a common item between the two instruments. This difference will be explored further in the section below. Learner gender does not seem to be a significant predictor of learner performance in the average score, but boys seem to have performed worse in the letter sound recognition. A measure of socio-economic status was included to control for household wealth. This indicator does not seem to explain as much of the variation in test scores as expected, but still shows that household wealth remains a predictor of learner outcomes even at the start of Grade 1. A district dummy was included which shows a slight difference in the overall score, but no difference in the letter sound recognition. However, overall the R-squared values are very low, which shows that we are not able to explain much of the variation in

general. Given the low correlations between the sub-tasks, it seems that the information gained from the assessments might be quite noisy.

Table 17: Learner Characteristics correlated with baseline performance

	Average Score (SD)					Letter sounds (Letters correct)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Learner age: 5 years or younger	-0.35*** (0.048)				-0.36*** (0.047)	-1.58*** (0.255)				-1.77*** (0.249)
Learner age: 7 years	0.25*** (0.044)				0.26*** (0.046)	0.41 (0.256)				0.72*** (0.246)
Learner age: Older than 7 years	0.06 (0.104)				0.09 (0.109)	-0.55 (0.464)				-0.01 (0.477)
Test Language: Zulu		0.06 (0.07)			-0.14 (0.091)		-1.16** (0.448)			-1.69*** (0.634)
Gender: Boy			-0.04 (0.035)		-0.04 (0.033)			-0.74*** (0.171)		-0.74*** (0.170)
SES				-0.05*** (0.018)	-0.05*** (0.016)				-0.2* (0.107)	-0.18* (0.100)
SES-squared				0.01 (0.013)	0.01 (0.013)				-0.11 (0.08)	-0.06 (0.074)
District dummy					0.16* (0.079)					0.33 (0.575)
N	3327	3327	3327	3281	3281	3327	3327	3327	3281	3281
Adjusted R ²	0.035	0.001	0.000	0.004	0.044	0.018	0.011	0.005	0.002	0.043

Note: Each column represents a separate regression on learner age, gender, home language, socio-economic background and district dummies. Standard errors are clustered at school level. * for p<.1; ** for p<.05; *** for p<.01

To further explore any potential language effects, the same regressions were run for each of the sub-tasks. Table 17 reports only the coefficients for the language and gender dummy, although were included for learner age, care-giver age, care-giver education level, literacy practices in the home relating to reading to the child and speaking English to the learner, socio-economic status and district effects. The results show that, even after controlling for these factors, six of the nine sub-tasks have significant language differences. These differences, however, are not necessarily favourable to a single language. SiSwati learners performed better than the isiZulu learners in phoneme isolation, letter sound recognition and English vocabulary, whereas isiZulu learners performed better in word recall, words read correctly and sentence reading.

As explained previously, the versioning of the instrument was conducted in such a way as to ensure that the exact same word/ sound was used where possible. This means that the letter sound recognition task is directly comparable between the two instruments. With regard to the phoneme isolation task, there was only one difference with the isiZulu assessment had the word “ceba” and the SiSwati assessment the word “liso”. The SiSwati version therefore had one easier item. Figure 21, indicates the percentage of learners who managed to successfully identify the initial sounds for each of the words. It is clear from this figure that a significantly

higher proportion of learners managed to successfully identify the initial sounds of the common items as well, and that the one easier item is not the only item driving the language difference. The significant difference on the English sub-task is specifically interesting as well, since this difference is apparent, even after controlling for the frequency of English spoken to the learner by the care-giver.

Although the language difference in the word reading and sentence reading tasks are significant, the performance in these two tasks are too low to make any substantial claims. When considering the word recall tasks, it is evident that the SiSwati version is linguistically slightly more difficult, which does not allow a direct comparison of the items.

Gender effects are also visible in four of the sub-tasks, with boys performing better than girls in the naming of the animals, but girls performing better in non-word recall, phoneme isolation and letter sound recognition. This is concurrent to what has been found in EGRS I for Grade 1 learners in Setswana.

Table 18: Learner Characteristics correlated with learner performance per sub-task

	Naming Animals	Word Recall	Non- word Recall	Phoneme Isolation	Compre- hension	Letter Sounds	Words Correct	Sentence Reading	Visual Perception	English Vocabulary
Test Language:	-0.2	0.49***	0.02	-0.44**	-0.02	-1.63**	0.72***	-0.05*	0.19	-0.32***
Zulu	(0.188)	(0.184)	(0.107)	(0.17)	(0.077)	(0.663)	(0.186)	(0.028)	(0.159)	(0.098)
Sex: Boy	0.65***	0.04	-0.08*	-0.18***	0.02	-0.66***	-0.05	0	-0.03	0.01
	(0.087)	(0.098)	(0.045)	(0.066)	(0.043)	(0.186)	(0.044)	(0.015)	(0.062)	(0.039)
N	3281	3281	3281	3281	3281	3281	3281	3281	3281	3281
Adjusted R ²	0.115	0.01	0.011	0.033	0.013	0.049	0.055	0.005	0.028	0.057

Note: Each column represents a separate regression on learner age, socio-economic background and district dummies. Standard errors are clustered at school level. * for p<.05; ** for p<.01; *** for p<.001.

Figure 21: Language Difference in Phoneme Isolation

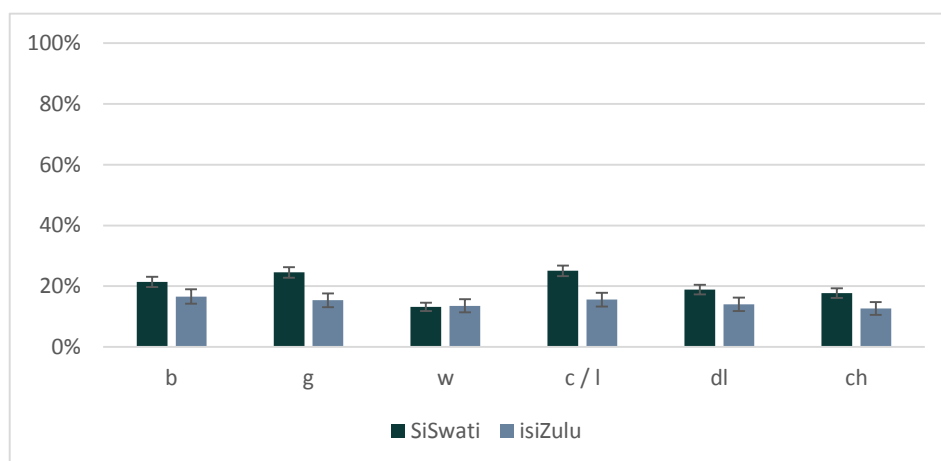


Table 18 reports the regression results of guardian or parent characteristics on learner performance. As before, each column represents a different regression, with the final column including all the variables included. The results show that learner's home characteristics play a significant role in predicting learner performance. The age of the learner's mother does not

seem to have any predictive value, whereas guardian's education and reading practices at home are positively correlated to learner performance. Learners from households where the guardian did not complete matric performed poorer than learners from households where the guardian did complete matric. Similarly, learners from households where they are read to every day performed significantly better than learners where reading happened less regularly (see Figure 21). For the English vocabulary task, the most significant predictor of learner performance was whether the guardian speaks English to the learner on a regular basis (see Figure 22).

Table 19: Learner background factors correlated with baseline performance

	Average Score (SD)					English Vocabulary				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Mother/ care-giver age	0.00 (0.002)				0.00 (0.002)	0.00 (0.011)				0.01 (0.013)
Did not complete matric		-0.14*** (0.042)			-0.12** (0.047)		-0.36 (0.224)			-0.25 (0.243)
Enrolled in/ completed any further training		-0.05 (0.062)			-0.02 (0.065)		-0.04 (0.360)			0.03 (0.382)
Read to child every day			0.09** (0.042)		0.07 (0.046)			0.25 (0.235)		0.19 (0.256)
Speaks English to the learner regularly			0.14** (0.067)		0.12 (0.075)			0.25 (0.317)		-0.09 (0.355)
Belief: Responsible for reading				0.10** (0.044)	0.07 (0.049)				0.40 (0.292)	0.29 (0.319)
Test-language: isiZulu					-0.02 (0.096)					-1.23* (0.670)
N	3070	3047	3192	3140	2682	3070	3047	3192	3140	2682
R-Squared	0.000	0.004	0.004	0.002	0.014	0.000	0.001	0.001	0.001	0.013

Note: Each column represents a separate regression on intervention dummies and stratification dummies. The final model controls for district effects and learner socio-economic status. Standard errors are clustered at school level. * for $p < .1$; ** for $p < .05$; *** for $p < .01$

Figure 23: Reading at home and learner performance

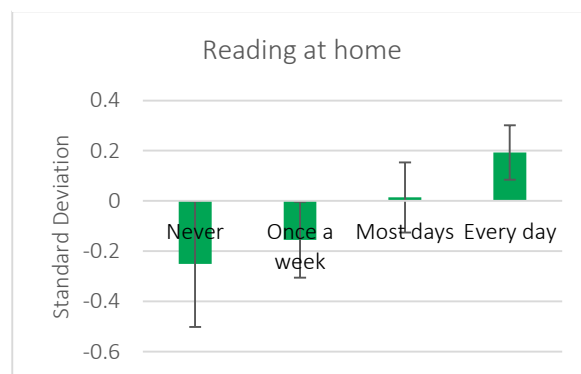
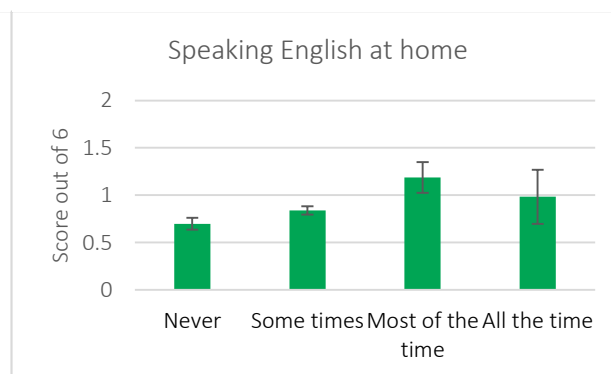


Figure 22: Speaking English at home and learner performance



Since the baseline assessment was conducted at the start of Grade 1, it is not expected that the school and teacher characteristics will have any direct correlation with learner performance, except for the aspects which might be a reflection of the community the learner is from. Various factors which provide one with an indication of the general wealth and education levels in the community were included in the regressions, but none of these factors predicted any variation in the learner performance.

Table 20: Community level factors

	Average Score (SD)			Letter sounds (Letters per minute)		
	(1)	(2)	(3)	(1)	(2)	(3)
Quintile 2	-0.01 (0.131)		0.15 (0.143)	-0.61		-1.07** (0.486)
Quintile 3	-0.03 (0.153)		0.19 (0.168)	-0.29		-1.25** (0.619)
Remote rural area	0.05 (0.124)		0.04 (0.145)	-0.18		0.01 (0.509)
Principal perception: At least 50% of parents have matric		0.05 (0.116)	0.11 (0.127)		0.15 (0.423)	-0.1 (0.426)
Principal perception: 20% - 40% of parents employed		-0.28* (0.142)	-0.26* (0.158)		0.88* (0.488)	0.71 (0.495)
Principal perception: 40% - 60% of parents employed		-0.27* (0.163)	-0.23 (0.189)		0.83 (0.584)	0.79 (0.607)
Principal perception: > 60% of parents employed		-0.13 (0.182)	-0.05 (0.182)		0.53 (1.126)	0.29 (1.141)
Gr 1 class enrolment			0.01 (0.006)			-0.03 (0.017)
Multi-grade school			-0.2 (0.216)			-0.6 (0.841)
LoLT: isiZulu			0.12 (0.191)			-1.24* (0.634)
District Dummy			0.31* (0.182)			-0.54 (0.615)
N	3,271	3,271	2,911	3,327	3,327	2,963
R-Squared	0.000	0.003	0.013	0.003	0.006	0.033

Note: Each column represents a separate regression on intervention dummies and stratification dummies. Standard errors are clustered at school level. * for $p < .1$; ** for $p < .05$; *** for $p < .01$

5. Sensitivity Checks

During the first term of implementing the interventions, two issues emerged which might affect the sample balance. Firstly, one school was dropped from intervention 2 due to a school consistently failing to send the Grade 1 teacher to the initial teacher training. Secondly, in inviting the schools to the initial teacher training, confusion arose regarding two schools with the same name. This has meant that the incorrect school attended the teacher training. With regards to the first issue, the project management team decided to still include the school in the data collection for the ensuing waves of data collection, but that the school will be considered as a non-compliant school. With regards to the second issue, the mistake was only recognised later in the study after implementation was well underway, and it was therefore decided to continue with the implementation as has been done for the first part of the year. The project management team determined that the mix-up was merely an administrative

mistake and therefore still consider the school allocation to be random. The concern however, lies with the impact that this will have on future analysis. Both schools are in the Gert Sibande District, both schools are Quintile 1 schools and both schools were classified as low performers for the stratification. The original school was in stratum 6 (large Quintile 1 schools who are weaker performance), whereas the new intervention school is in stratum 2 (medium-sized schools from all quintiles who are weaker performers).

Sensitivity analyses will be conducted in all future reports to gauge the effect that this new allocation have on the final results. Table 20 below shows the balance test results once the intervention status of the two schools were swopped.⁸ It is evident that this change does not affect the balance of the sample significantly, with only a minor additional imbalance in the visual perception task. This difference, however, does not influence the overall balance of the sample. A further test was done by dropping both schools from the sample, but this did not have any significant difference on the sample balance.

Table 21: Robustness check - Re-allocating intervention status

	Naming Animals	Word Recall	Non- word Recall	Phoneme Isolation	Compre- hension	Letter Sounds	Words Correct	Sentence Reading	Visual Perception	English Vocabulary
Paper- Based	0.15 (0.154)	-0.01 (0.118)	-0.02 (0.068)	-0.06 (0.12)	-0.01 (0.064)	-0.19 (0.486)	-0.06 (0.137)	-0.03* (0.016)	0.20* (0.12)	-0.02 (0.067)
Tablet- Based	0.39*** (0.149)	0.12 (0.118)	0.03 (0.095)	0.04 (0.13)	0.1 (0.063)	0.06 (0.477)	0.13 (0.164)	-0.02 (0.016)	0.21 (0.125)	0.01 (0.057)
Control Mean	7.155	9.981	4.208	1.129	2.179	4.652	0.387	0.051	1.46	0.836
N	3327	3327	3327	3327	3327	3327	3327	3327	3327	3327
Paper= Tablet: p-value	0.154	0.259	0.557	0.448	0.101	0.649	0.26	0.507	0.965	0.594

⁸ The stratum allocation of the two schools were also swopped around. This decision essentially makes the assumption that the allocation of these schools to their new intervention status was random, and could have been based on the stratum they were allocated to. This ensures that the strata also remain equally sized.

6. USAID indicators

Table 22: Specific USAID Indicators

			FY October 2015 - September 2016				FY October 2016 - September 2017			
	Indicator no	Indicators	Target	Actual	Male	Female	Target	Actual	Male	Female
Specific Indicators	ES. 1-1	Proportion of students who, by the end of two grades of primary schooling, demonstrate that they can read and understand the meaning of grade level text.	0%	0%	0%	0%	0%	0%	0%	0%
	ES. 1-1a;b	Percent learners	0%	0%	0%	0%	0%	0%	0%	0%
	ES. 1-1c;d	Numerator	0	0	0	0	0	0	0	0
	ES. 1-1e;f	Denominator	0	0	0	0	0	0	0	0
	ES. 1-5	Number of learners reached in reading programs at the primary level	0	0	0	0	7000	7600	4114	3486
	ES. 1-7	Number of primary school educators who complete professional development activities on implementing evidence-based reading instruction with USG assistance	0	0	0	0	212	188	0	188
	ES. 1-11	Number of primary school classrooms that receive a complete set of essential reading instructional materials with USG assistance	0	0	0	0	212	188	0	188

Table 23: Custom USAID Indicators

	Indicator no	Indicators	FY October 2015 - September 2016				FY October 2016 - September 2017			
			Target	Actual	Male	Female	Target	Actual	Male	Female
Custom Indicators	2.1.	Proportion of teachers that attended each compulsory training session	0	0	0	0	95%	94%	0%	100%
	2.2.	Number of principals/HOD's that attended each training session	0	0	0	0	50	27	1	26
	2.3.	Number teachers who attended needs driven workshops	0	0	0	0	50	66	0	66
	2.4.	Average number of interactions between the virtual coach and a teacher	0	0	0	0	1 per week	1 per week	0	93
	2.5.	Average number of visits a teacher receives from a reading coach.	0	0	0	0	1 per month	0.75 per month	0	93
	2.6.	Comparison of baseline and endline scores	0	0	0	0	0	0	0	0
	3.1.	Number of learner background questionnaires completed, as well as actual data	0	0	0	0	3,600	3,327	1,801	1,526
	3.2.	Number of teacher questionnaires completed, as well as actual data	0	0	0	0	360	320	0	320
	3.3.	Number of principal questionnaires completed, as well as actual data	0	0	0	0	180	180	98	82
	4.1.	Number and details of key stakeholders involved in planning	10	0	0	0	5	4	2	2
	4.2.	Number and details of research dissemination sessions with provinces and schools	0	0	0	0	5	4	.	.
	4.3.	Number and details of reports, journal articles and conferences	0	0	0	0	0	0	0	0

7. Next Steps in the project

Interventions continued throughout 2017, with the initial training having been conducted just after the conclusion of the baseline data collection. Teachers further attended clustered workshop training for one day at the start of each term, and the reading coaches provided on-going support throughout the year. The midline data collection is taking place from 23 October to 10 November 2017, and will provide the data necessary to determine the impact of the interventions after one year of implementation. The interventions are set to continue to Grade 2 teachers in 2018, and the same learners will be tested at the end of Grade 2 in 2018.

References

- Adams, M., 1990. *Beginning to Read: Thinking and Learning about Print*. Cambridge: MIT Press.
- Aploon-Zokufa, K., 2013. Locating the difference: A comparison of pedagogic strategies in high and low performing schools. *South African Journal of Childhood Education*, 3(2), pp. 112-130.
- Boakye, N., 2012. *A socio-affective approach to improving students' reading comprehension abilities*. Pretoria: Faculty of Humanities. University of Pretoria.
- Coe, R., Aloisi, C., Higgins, S. & Major, L., 2014. *What makes great teaching? Review of the underpinning research*. London: Sutton Trust.
- Diemer, M., Van der Merwe, K. & De Vos, M., 2015. The development of phonological awareness literacy measures for isiXhosa. *South African Linguistics and Applied Language Studies*, 33(3), pp. 325-341.
- DPME, forthcoming. *CAPS Implementation Evaluation*, Pretoria: Department of Planning, Monitoring and Evaluation.
- Duke, N., 2000. For the Rich It's Richer: Print Experiences and Environments Offered to Children in Very Low- and Very High-Socioeconomic Status First-Grade Classrooms. *American Educational Research Journal*, 37(2), pp. 441-478.
- Fleisch, B. & Christie, P., 2004. Structural Change, Leadership and School Effectiveness/Improvement: Perspectives from South Africa. *Discourse: Studies in the Cultural Politics of Education*, 25(1), pp. 95-112.
- Fleisch, B. & Schoer, V., 2014. Large-scale instructional reform in the Global South: insights from the mid-point evaluation of the Gauteng Primary Language and Mathematics Strategy. *South African Journal of Education*, 34(3), pp. 1-11.
- Fleisch, B. & Schoer, V., 2014. Large-scale instructional reform in the Global South: insights from the mid-point evaluation of the Gauteng Primary Language and Mathematics Strategy.. *South African Journal of Education*, Volume 3, pp. 1-34.
- Fleisch, B. T. S., Schoer, V. & Mabogoane, T., 2017. Failing to catch up in reading in the middle years: The findings of the impact evaluation of the Reading Catch-Up Programme in South Africa. *International Journal of Educational Development*, 53(1), pp. 36-47.
- Gathercole, S., 2006. Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, Volume 27, pp. 513-543.
- Hill, H., Charalambos, Y. & Kraft, M., 2012. When Rater Reliability Is Not Enough: Teacher Observation Systems and a Case for the Generalizability Study. *Educational Researcher*, pp. 56-64.
- Hoadley, 2012. *What do we know about teaching and learning in South African primary schools?*. Cape Town, Towards Carnegie III Conference.
- Hoadley, U., 2012. What do we know about teaching and learning in South African primary schools?. *Education as Change*, 16(2), pp. 187-202.
- Hoadley, U. & Galant, J., 2016. Specialisation and school organisation: Investigating pedagogic culture.. *British Journal of Sociology and Education*, 37(8), pp. 1187-1210.

- Howie, S. et al., 2012. *PIRLS 2011 South African Children's reading literacy achievement: Summary Report*, Pretoria: Centre for Evaluation and Assessment, University of Pretoria.
- Macdonald, C., 2002. Are children still swimming up the waterfall? A look at literacy development in the new curriculum. *Language Matters*, 33(1), pp. 111-141.
- Macdonald, C., 2006. The properties of mediated action in three different literacy contexts in South Africa. *Theory and Psychology*, 16(1), pp. 51-80.
- Matjila, D. & Pretorius, E., 2004. Bilingual and Biliterate? An exploratory study of Grade 8 reading skills in Setswana and English.. *Per Linguam*, 20(1), pp. 1-21.
- McDonald, C. et al., 2009. The ISI Classroom Observation System: Examining the Literacy Instruction Provided to Individual Students. *Educational Researcher*, 38(2), pp. 85-99.
- NEEDU, 2013. *National Report 12. The state of literacy teaching and learning in the Foundation Phase*, Pretoria: Department of Basic Education .
- 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), pp. 39-45.
- 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), pp. 11-21.
- Piper, B., Zuilkowski, S. & Mugenda, A., 2014. Improving reading outcomes in Kenya: First-year effects of the PRIMR Initiative.. *International Journal of Educational Development*, Volume 37, pp. 11-21.
- Pretorius, E. & Currin, S., 2010. Do the rich get richer and the poor poorer? The effects of an intervention programme on reading in the home and schools language in a high poverty multilingual context.. *International Journal of Educational Development*, 30(1), pp. 67-76.
- Pretorius, E. & Lephala, M., 2011. Reading Comprehension in High-Poverty Schools: How should it be taught and how well does it work?. *Per Linguam*, 27(2), pp. 1-24.
- Pretorius, E. & Mokhwesana, M., 2009. Putting reading in Northern Sotho on track in the early years: Changing resources, expectations, and practices in a high poverty school. *South African Journal of African Languages*, 1(1), pp. 54-73.
- Raudenbush, S., 2005. Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, 34(5), pp. 25-31.
- Reeves, C. et al., 2008. *Evaluation of literacy teaching in primary schools in Limpopo*, Polokwane: Limpopo Department of Education.
- Snijlsteit, B. et al., 2016. *The impact of education programmes on learning and school participation in low- and middle-income countries: a systematic review summary report, 3ie Systematic Review Summary 7*, London: International Initiative for Impact Evaluation (3ie).
- Spaull, N., 2016. *Learning to Read and Reading to Learn*. Policy Brief. Research on Socioeconomic Policy (RESEP): (Online). Available: www.resep.sun.ac.za.
- Taylor, N., 2007. Equity, efficiency and the development of South African schools. In: T. Townsend, ed. *International Handbook of School Effectiveness and Improvement*. Springer International Handbooks on Education. Dordrecht: Springer.

Taylor, N., Van der Berg, S. & Mabogoane, T., 2013. *Creating Effective Schools*. Cape Town: Pearson.

Taylor, S., 2011. Uncovering indicators for effective school management in South Africa using the National School Effectiveness Study. *Stellenbosch Working Paper Series: 10/11*.

Taylor, S., 2013. Modelling educational achievement. In: N. Taylor, S. Van der Berg & T. Mabogoane, eds. *Creating effective schools*. Cape Town : Pearson Education South Africa, pp. 63-101.

Taylor, S. & Watson, P., 2015. The impact of the study guides on "matric" performance: Evidence from a randomised experiment. *Stellenbosch Economic Working Papers: 13/15*.

Van der Berg, S., 2009. Fiscal incidence of social spending in South Africa. *Stellenbosch Working Paper Series: 10/09*.