Improving Early-Grade Reading in South Africa

Pre-analysis plan (PAP)

28 October 2016

Jacobus Cilliers and Stephen Taylor

BACKGROUND

This project is a Randomised Controlled Trial (RCT) to evaluate the impact of three alternative treatments all aimed at improving early grade reading in the home language in the South African school system.

The primary implementing partner is the South African government, in particular the Department of Basic Education. A key role is also being played by the North West provincial education department, which is contributing financially and is championing the project within the schools.

A service provider has been appointed to run the three interventions on behalf of the DBE for the purposes of this impact evaluation. The service provider is an organisation called "Class Act", which is highly involved in partnerships with government to run literacy interventions. For example, "Class Act" was a service provider in the Gauteng Province's implementation of the Gauteng Primary Literacy and Maths Strategy (GPLMS) over the last few years.

The evaluation side of the project is being supervised by the Research Team while the data collection and capturing is being managed by South Africa's Human Sciences Research Council (HSRC).

The Research Team consists of academics from the University of Stellenbosch, the University of the Witwatersrand, Georgetown University and the Human Sciences Research Council.

DESCRIPTION OF INTERVENTIONS

This study evaluates three different interventions, all aimed at improving early-grade reading in the home language, which in the case of the North West province is Setswana. All three interventions work with children entering Grade 1 at the start of 2015 over a two-year period (thus working with grade 2 learners in 2016).

Treatment 1: Training, scripted lessons, graded readers.

Treatments 1 and 2 aim to apply the same set of instructional practices in the teaching of home language literacy in grade 1 and 2 classrooms. Both treatments therefore provide teachers with clearly scripted lesson plans, which are aligned to the curriculum as specified in the Curriculum and Assessment Policy Statements (CAPS) for home language literacy in the Foundation Phase. The lesson plans incorporate the use of learning support materials including the government-provided workbooks as well as certain additional materials (graded reading booklets, flash cards, posters, etc.), which are provided through the EGRS. The graded reading booklets provide a key resource for the teacher to use in group-guided reading and individual work so as to facilitate reading practice at an appropriate pace and sequence of progression.

Treatment 1 trains the teachers on how to use the lesson plans and accompanying materials through central training sessions, each lasting 2 days, and occurring twice yearly. The first session was conducted in January 2015 and the second occurred in July 2015. Similar sessions are scheduled for 2016.

Treatment 2: Reading Coaches, scripted lessons, graded readers.

Exactly the same set of instructional materials (scripted lesson plans, graded reading booklets and other materials) is provided to Treatment 2 schools. However, instead of central training sessions, ongoing support to teachers consisting of regular (monthly) on-school coaching from specialist "reading coaches" is provided. In addition to these on-site visits, there are occasional meetings with the coach and a small cluster of nearby Treatment 2 schools. The evaluation of treatments 1 and 2 should thus shed light on whether the fairly prescriptive instructional regime has the ability to improve reading acquisition and whether the mode of teacher support is important in mediating effectiveness.

Treatment 3: Parental involvement

Treatment 3 is designed to promote parental involvement to support their children's reading progress. At each of the 50 schools in this treatment arm a Community Reading Coach (CRC) was recruited. The CRC was identified through communication with the school principal who recommended a suitably qualified but available person in the community. The CRCs attend a 1-day training session facilitated by the service provider (Class Act) at the start of each school term (quarterly). The CRCs are trained to deliver weekly training sessions for grade 1 parents at their respective schools. A total of 30 sessions is scheduled for each year covering a total of 10 topics. Each topic has 3 sessions where the topic is the same but the activities of the session differ. Thus a parent can attend roughly 1 in 3 sessions and still be exposed to all topics, while parents who attend more regularly can still enjoy fresh activities. For their services, CRCs are paid a stipend of R400 per month (about $35).

The topics covered in these sessions include the importance of learning to read for later educational and labour market success, training on how to support their child's reading at home and the provision of low-cost materials and reading games to use at home.

THEORY OF CHANGE

Reading acquisition – vocabulary and decoding

All our interventions are grounded in the educational theory of how reading acquisition occurs. As a starting point a pupil needs to develop vocabulary and master decoding (letter recognition and phonetic awareness). Only then can a pupil progress towards word recognition and, with sufficient practice, eventually fluency. Letter recognition and phonemic awareness do not come naturally and need to be taught: they are mastered through systematic teaching and consistent practice. Practice occurs both at school and at home.

In order to learn the basics of decoding, a child requires a teacher who is present, capable and motivated to deliver systematic reading instruction. In order for decoding to become fluent a child requires suitable graded materials and the discipline (perhaps imposed) and opportunities to practice a lot, both at school and home. The interventions tested in this study address these needs in various ways.

Interventions (1) and (2)

Scripted lessons provide a structure to assure systematic practice and learning based on sound pedagogical theory. In addition, they require no additional lesson preparation from teachers, so make it easy for teachers can switch to a more productive teaching practice with limited additional effort. Furthermore, scripted lesson plans free up teachers' time, because they no longer need to allocate time to preparation. This could improve reading acquisition if teachers allocate this time to productive teaching activities, rather than leisure or unproductive teaching activities.

The accompanying graded reading materials provide ample material for pupils to practice decoding and reading at their level of development: the opportunity to learn 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 pupils' reading proficiency in order to assign pupils to the appropriate graded readers and smaller reading groups, based on ability. The group-guided reading also provides the teacher opportunities to provide individualized attention.

A growing body of evidence has shown that the curriculum is too ambitious in many developing countries and children are being left behind. This is plausibly also the case in South Africa, since the majority of grade 5 pupils still cannot read for meaning. Interventions that free up resources for teachers (or assistant teachers) to teach to the level of the child

rather than the curriculum can have a large impact on learning (Banerjee et al. 2007, Duo & Kiessel 2014, Duflo et al. 2011).

In addition to the above, the reading coach intervention provides more intensive training to improve teacher capacity. 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 reading coaches could also improve teacher motivation as they are frequently monitored, provided with much-needed additional support, and can also find inspiration from watching an excellent example provided occasionally by coaches.

Intervention 3:

Parents pay a critical component to learning to read, as it requires continuous practice, both at school and at home. For parents to be *willing* to play this role they need to appreciate (i) the importance of reading; and (ii) and their child is most likely not learning enough at school and requires additional support. This is the purpose of the information. For parents to be *able* to play this role, they need to understand the necessary steps in learning to read and also have appropriate material to practice reading with their child. This is the purpose of the training and additional practice material.

METHODS OVERVIEW

DATA SOURCES

The baseline and midline data collection had already taken place in February and November 2016. Endline data collection is planned for November 2016.

We will visit the same 230 schools and measure the following:

- Pupil assessment in reading proficiency.
- Basic teacher assessment in reading/comprehension proficiency.
- School principal survey
- Teacher survey
- Facility and document review
- Parent survey

In addition we will perform classroom observation and more detailed document review in a subset of 60 schools – 20 schools in treatments 1, 2 and the control group respectively.

IDENTIFICATION AND EMPIRICAL STRATEGY

Schools are randomly assigned.

Our main estimating equation will be:

$$y_{isb1} = \beta_0 + \beta_1 T + X'_{isb0}\Gamma + \rho_b + \varepsilon_{isb1}$$

Where y_{isb1} is the is the outcome indicator of interest (more below) for pupil i in school s and strata b, T is the treatment dummy which is equal to one in either of the treatment arms, ρ_b refers to strata fixed effects, X'_{isb0} is a vector of baseline controls, and ε_{isb1} is the error term clustered at the school level. To estimate the respective impacts of the three interventions, we restrict the sample to the control schools and the schools from the relevant treatment group.

We will control separately for each domain of reading proficiency collected at baseline: vocabulary, letter recognition, working memory, phonological awareness, word recognition, words read, and sentence comprehension. We will control for each domain separately, rather than create an aggregate index of learning proficiency in order to increase statistical power. To further increase statistical power and account for any incidental differences that may exist between treatment

groups, we will control for individual and community-level characteristics which are highly correlated with y_{isb1} or were imbalanced at baseline.

TREATMENT ASSIGNMENT AND SAMPLE SELECTION

Through a process of elimination we developed a sampling frame of 230 eligible schools. Beginning with 458 primary schools registered in 2014 administrative data in the districts of Dr Kenneth Kaunda and Ngaka Modiri Molema we started by excluding relatively affluent schools (those in quintiles 4 and 5). Next, we excluded schools in which the language of instruction in the Foundation Phase was not Setswana. We excluded schools which were missing in the 2014 ANA dataset. We also excluded 8 schools that had already been selected for the purposes of piloting of instruments through the course of this project. We further excluded particularly small schools (fewer than 20 grade 1 enrolments) since many of these schools would practice multi-grade teaching rendering the scripted lesson plans less appropriate. We also excluded particularly large schools (more than 180 grade 1 enrolments) to limit intervention costs. Three more schools were excluded after the North West PED checked our list of schools and found specific problems with these schools (e.g. the school had been closed down, or a particular conflict around school management was occurring in a school). After all of these exclusions 235 eligible schools remained. Using a random number generator, we then excluded 5 schools, which we retained as possible replacement schools. Thus we obtained the sampling frame of 230 schools.

To increase power and assure balance between treatment arms, we performed stratified randomization. We created 10 strata of 23 similar schools based on school size, socio-economic status, and previous performance in the Annual National Assessments. Within each stratum, we then randomly assigned 5 schools to each treatment group and 8 to the control group. Thus we randomly assigned 50 schools to each treatment and 80 to the control. Given that we 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 a treatment group with the control group and a minimum effect size of 0.23 standard deviations when comparing two treatment groups. These calculations assume a 95% confidence interval, an alpha value of 0.8, an intra-class correlation coefficient (rho) of 0.3 and a correlation between pre- and post-test scores of 0.7.

HYPOTHESES UNDER INVESTIGATION

This section outlines the main and intermediate outcomes we want to measure, as well as heterogeneous treatment effects of interest. In brackets we indicate the source of the data as well as question number if it is based on a survey. If an outcome is based on multiple indicators, it will be aggregated using principal component analysis.

[*SPQ = School Principal Questionnaire; TRQ = Teacher Questionnaire; PRQ = Parent Questionnaire; CL = Classroom Observation*]

MAIN OUTCOMES

Our main research questions are:

1. Did treatment one (one-off training) improve learner reading proficiency?
2. Did treatment two (reading coaches) improve learner reading proficiency?
3. Did treatment three (parental involvement) improve learner reading proficiency?
4. Did the impact on reading proficiency differ between the three treatment arms?
5. Which one is most cost-effective?

For reading proficiency we will measure all the intermediate steps towards comprehension: letter recognition, phonetic awareness, word recognition, fluency, and reading with comprehension. These tests will be adapted from standard tests that have already been developed for the Setswana language, such as the Early-Grade-Reading-Assessment (EGRA). As our main outcome indicator, we will also construct an aggregate indicator of learning proficiency, using principal component analysis.

INTERMEDIATE OUTCOMES

In order to test for mechanisms based on our theory of change, we ask the following:

TEACHERS

Did treatment arms one and two change **teaching practice**?

1. Provide regular individualized assessment [*TRQ – 5(e)*]
2. Require student to read out loud [*TRQ – 5(f)*]
3. Stream by ability [TRQ – 5(i)]
4. Assess reading ability [TRQ - *Document review* – 9]
5. The completion of more writing exercises [TRQ - *Document review* – 13-14]
6. The assessment of more writing exercises [TRQ *Document review* – 15]
7. Increase the availability of home language text in classrooms [TRQ - *Classroom observation* – 8b]
8. Improved knowledge and practice of appropriate instructional practices prescribed in the official curriculum [*TRQ – 5(j)(k)*]

Did treatment arms one, two or three cause higher **teacher effort**?

1. Presence [SPQ *Document Inspection* – 5; *TRQ 2(c)*]
2. Lesson Preparation *during* class hours. [*TRQ – 2(d)*]
3. Use of the national workbook [*TRQ document inspection – Q11*]

Based on the **classroom observation** toolkit, did treatments one or two change any the following teaching practice?

1. Listening and speaking
    a. Pupils answer in unison [CL – Q34]
    b. Same pupils always answer the question [CL – Q35-36,Q38]
    c. Pupils understand tasks [CL – Q37]
    d. Provide individualized attention [CL - Q25]
2. Use of reading material
    a. Number of pupils who handle books [CL – Q42]
    b. Number of pupils who read [CL – Q43]
3. Vocabulary development [CL – Q59-61]
4. Opportunities to write [CL – Q62,Q67]
5. Opportunities to read [CL – Q72-74
6. Tracking
    a. Writing difficulty based on ability [CL - Q68]
    b. Reading groups of different ability [CL – Q75]
7. Teaching to the level of the child
    a. Ask questions that are too difficult/easy [CL – Q79]
    b. Reading material too difficult/easy [CL – Q79]
    c. Writing activities too difficult/easy [CL – Q79]
8. Assessment
    a. Teacher corrects pupil [CL – Q39 – 40]
    b. Written work is marked [CL – Q82]
    c. In-class reading assessment [CL – Q83]
    d. In-class phonics assessment [CL – Q84]

PARENTS

Did treatment arm three lead to a change in **parents' beliefs and aspirations**? [*parent survey*]

1. Personal sense of efficacy and responsibility. [PRQ – 12]
2. Belief in child's reading ability [*PRQ – 13*]

Did treatment arm three lead to changed **parent behavior**?

1. Read to child [PRQ – 6 and 8]
2. Play educational games with child [PRQ – 9 and 10]

3. Attend PTAs/SGB [*PRQ – 11*]
4. Check child's homework [PRQ – 5 and 16]
5. Provide opportunity for child to do homework [PRQ – 16]
6. Provides structure and regular hours (goes to bed early, etc.) [PRQ – 17]

PUPILS

1. Pupil attendance [PRQ – 14]
2. Pupil punctuality [TRQ 17

HETEROGENEOUS IMPACTS

Note that the theory of change outlined above predicts many heterogeneous treatment impacts.

At the pupil level, we could expect two opposing heterogeneous treatment impacts of treatments one and two based on baseline pupil reading proficiency. The scripted lesson plans require streaming by ability *within* the same classroom and provides opportunity for individualized attention and could benefit children who have otherwise been left behind. However at the same time the scripted lesson plans are aligned to the national curriculum, which prescribes an ambitious pace in the South African context. The worst-performing pupils might actually benefit less, if the teachers who follow the scripted lesson plans now progress at too fast a pace. Furthermore, boys/girls might benefit more/less from the individualized attention. Finally, the emphasis on individualized attention and tracking means that pupils might benefit more from the scripted lesson plans when the class size is large, and the worse-performing pupils in particular will benefit more.

At the teacher level, we should expect that the success of the interventions one and two depend on teacher motivation, prior levels of effort, and ability. The scripted lesson plans will only be applied by teachers who have a sufficient level of intrinsic motivation. Related, teachers who have a higher burden of lesson preparation are most likely to switch to scripted lesson plans, because they have most to gain from no longer test. Furthermore, teachers may need a sufficient baseline level of reading proficiency in order to effectively apply the scripted lesson plans. On the other hand, the scripted lesson plans might be too restrictive for exceptional teachers who are effective at adjusting their instruction to the needs of the classroom and ability of the pupils.

At the community and school level, we could again expect opposing heterogeneous impacts based on parents' baseline characteristics. Intervention three could be less effective with very educated and involved parents because they are already making all the necessary investments in their child's learning progress, or it could be more effective if a baseline level of parent support is required for the interventions to work. Interventions one and two might be more effective in schools that already have a minimal baseline level of performance so that children can benefit from the changed practice, or it might be most beneficial in the worst-performing schools where the need for improved teaching practice is highest.

We will therefore examine the following possible interactions:

PUPIL LEVEL

1. Pupil gender [pupil baseline test]
2. Pupil age [pupil baseline test]
3. Pupil baseline performance [pupil baseline test]
4. The interaction between baseline performance and class size [pupil test and SPQ]

TEACHER LEVEL

1. Baseline teacher characteristics:
    a. Education level [TRQ – 3(a)]
    b. Performance in the teacher reading test [Teacher written test]
    c. Years of experience [TRQ – 3(c) and 3(e)]
    d. Own background reading [TRQ – 4]
2. Baseline teacher presence
3. Whether teacher spent time after class hours preparing lessons
4. Previous training opportunities [TRQ – 6(a) and 6(b)]

5. Whether the teacher tracks pupils [TRQ – 5(l)]
6. The extent to which teachers feel supported and recognized in their work [TRQ – 5l]
7. The sufficiency of the classroom infrastructure [TRQ – 8b]

SCHOOL-LEVEL

1. Baseline school average reading performance [administrative data]
2. Socio-economic background of school [administrative data]
3. School resources [SPQ: 4(a), 6, and 7]
4. School location [SPQ – 4(b)]
5. Pupil-teacher ratio [SPQ 2(c) and administrative data]
6. Overall educational background of the parents [SPQ – 3(f)]
7. Overall employment of parents [SPQ – 3(g)]

PARENT LEVEL

1. Identity of the parent/guardian: [PRQ - 2]
   E.g. single mother/grandmother; child-headed household.
2. Education level [PRQ – 4]
3. Baseline written ability of parent [PRQ – whether they can enter Q11]
4. Baseline involvement in child's learning [PRQ – 5,6 and 9; 12 and 13]

ADDITIONAL CORRECTIONS

CORRECTION FOR MULTIPLE HYPOTHESES

1. **Mean index:** To minimize risk of over-rejection of the null, we will construct a mean index of the different metrics of reading proficiency, using principal component analysis. However, since each metric relates to a different component in reading progression it is important to test each separately to identify blockages in reading progression.
2. **Multiple comparison corrections:** This will not be necessary as we only have one main outcome of interest: reading proficiency. Furthermore, since each treatment is conceptually different with a different theory of change and resource intensity, it does not make sense to make multiple-comparison corrections across treatment arms. Although we will be testing many different sub-group effects, the main impact of program will be evaluated on reading proficiency.
3. **Exclusion of multi-**grade schools: In a small number of schools there is multigrade teaching, e.g. grade 2 and 3 is combined in a single class. This limits the appropriateness of following the lesson plans provided. Therefore, as a robustness check, the main model will be re-run excluding those multi-grade schools.

ADDRESSING SURVEY ATTRITION

We will construct a dummy indicating whether a pupil attrited between baseline and midline (endline) and test if attrition was imbalanced (by regressing the attrition dummy on treatment status) and also if it was non-random (by regressing the attrition dummy on baseline test scores).

If attrition was found to be non-random and imbalanced, we can construct Lee Bounds – a conservative measure of the upper and lower bounds based on the most extreme sample selection - and also conduct Inverse Probability Weighting.

COST EFFECTIVENESS

In order to establish the cost effectiveness of the each treatment we calculate the standard deviations gained per US$100 spent on treatment. This allows us to make comparisons with other studies reported on in Kremer, Brannen and Glennerster (2013). We use the estimated treatment effect sizes from the main equation. We convert costs from Rand values to US dollars using the Rand-Dollar exchange rate as at the close of the South African markets on October 21 2015. The rate to be used is thus R13.40 to US$1.