

https://dspace.org/

The effects of teaching assistants on letter knowledge of foundation phase students in South Africa Evidence from exploratory trial

Brahm Fleisch; Volker Schoer

https://hdl.handle.net/10539/42104
SCOPUS

Downloaded from a DSpace repository. https://dspace.org/

International Journal of Educational Research Open 7 (2024) 100383

Contents lists available at ScienceDirect

International Journal of Educational Research Open

journal homepage: www.elsevier.com/locate/ijedro

The effects of teaching assistants on letter knowledge of foundation phase students in South Africa: Findings from an exploratory trial

Brahm Fleisch [a,*], Volker Schöer [b,c]

[a] Wits School of Education, University of the Witwatersrand, South Africa
[b] School of Economics and Finance, University of the Witwatersrand, South Africa
[c] School of Business and Economics, University of Tübingen, Germany

ARTICLE INFO

Keywords:
Letter knowledge
Teaching assistant
Early grade reading

ABSTRACT

There is increasing evidence that letter knowledge is a significant predictor of early grade reading achievements and that interventions designed to improve learning of this constrained skill can be effective. However, little is known about ways to enhance the teaching of letter knowledge in African languages and the degree to which teaching assistants can facilitate this process. This paper reports findings from an exploratory study designed to establish if teaching assistants can improve letter knowledge in African languages. A pre-post-trial with 249 pre-Grade R and Grade 1 students was conducted in two South African provinces with two distinctive African languages, isiXhosa and Sepedi. Our analysis shows significant positive correlations for both languages between students' exposure to teaching assistants and their improvements in the number of attempted letters as well as correct letters in a standardised test. Therefore, our findings suggest that teaching assistants, who are instructed to play letter games with foundation phase students to complement the teaching of letter knowledge, could be an effective, low-cost strategy to facilitate future improvements in reading.

1. Introduction

The problem of the learning crisis, particularly the failure of South African children to learn to read for meaning, is now well established (Spaull & Pretorius, 2019). There is an emerging body of research that is suggesting that the problem stems from very early in the school career with the failure to master the building block of reading, that is, letter names and letter sounds (Ardington, et al 2021; Nguma, 2021).

In one of their interviews, Nguma (2021) quotes a Grade 3 teacher: "In my Grade 3 class, I have students who have a challenge in independent reading. They can't read independently. They fail to recognise most letter sounds in a text. I think they did not master the letter sounds in Grade 1 and 2. (p. 203)"

How can we address this specific aspect of the wider learning crisis in South Africa? An opportunity has been presented through the implementation of the Presidential Youth Employment Initiative (PYEI) through the Department of Basic Education (DBE) that offered Teaching Assistant positions to unemployed youth in 2022 and 2023. Although the Teaching Assistant programme's (PYEI-DBE initiative[1]) main objective is to provide work opportunities to hundreds of thousands of young people who had been excluded from employment opportunities, little is known about if, and to what degree, teaching assistants could impact the learning crisis. For our context, two questions inform this pilot study. The first relates to the type and level of implementation of and enactment by the teaching assistants. The second relates to the potential impact of the intervention on the key foundation phase learning outcomes.

The evaluation of the study was structured around five steps. The first relates to the overall model that the intervention deploys. The second explores how the model is operationalised as a programme. This includes the scope and sequence of the intervention, tasks and activities, and budget. The third step involves an implementation evaluation of the extent to which the programme as designed gets implemented. For example, were the materials that were part of the design printed and given to the correct recipients? Did training take place, of what quality and what were the levels of attendance? The next step involves translating the programme into practice in the classroom. The evaluation must establish the extent to which TAs enacted the practice daily. More to the point, given the number of disruptions and early end of the school year, how many actual sessions on average occurred? Finally, the

* Corresponding author.
E-mail address: brahm.fleisch@wits.ac.za (B. Fleisch).

[1] See for details: https://www.stateofthenation.gov.za/assets/downloads/PES-Report-BEEI-spotlight.pdf

https://doi.org/10.1016/j.ijedro.2024.100383
Received 10 May 2024; Received in revised form 13 August 2024; Accepted 13 August 2024
Available online 13 September 2024
2666-3740/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

evaluation needs to rely on learning outcomes to provide insights into the impact of the intervention. Depending on the nature of the design, the impact evaluation may or may not be able to make policy-relevant claims. Considering that this study followed a simple pre-post research design instead of a counterfactual design, the findings cannot be interpreted causally.

The paper is divided into five sections. In the first section, we review the relevant literature on teaching letter knowledge. The second section provides insights into the background and features of the intervention model. The third section describes the evaluation design, examining both its limits and potential benefits for deriving policy insights. In this section, we also analyse levels of implementation and enactment. The fourth section describes the main findings. In this section, we explore pre- and post-results from student testing, comparing both the internal data gathered by the teaching assistants and letter knowledge assessments from an independent, external agency tasked with the data collection processes. We explore various aspects of student performance including average effects and effects of their performance distributions, as well as changes in the percentage of students at or above various DBE foundation phase learning outcome benchmarks. The final section contains the discussion and explores what if any implications for policy can be drawn from this pilot study.

2. Literature review

The international literature on the link between the failure to master letter names and letter sounds and the failure to learn to read is now well established (McBride-Chang et al, 1999; Hulme et al 2012). Noel Foulin (2005) review of the literature on the relationship between letter knowledge and reading suggest that there may be three mechanisms. Children who have strong letter knowledge are more likely to develop phonological processing of print, are better able to learn the letter-sound correspondences and are more likely to develop phonemic sensitivity skills. Treiman's (2013) review explores how children learn about letters, from the letters in their names to the links between letters and phonological skills that link the letter names to the letter sounds. Again, the emphasis is that the letter knowledge is the key to linking print and speech. She notes that although letter knowledge is not a sufficient condition for becoming a fluent reader (particularly in English), it is an important first step towards mastery of the writing system. Sunde et al., (2020), while acknowledging the centrality of mastering letter knowledge as a precondition for learning to read, address the question about the pace at which new letters should be introduced. The common practice in many countries that use the Roman alphabet is to introduce one letter per week, suggesting that children would master the entire alphabet in roughly half a school year by the middle of the Grade 1. Sunde et al. (2020) found that a faster pace of letter instruction can yield significantly better results for all reading outcome measures, with lower performing students benefiting more than higher performing students. Finally, although not specific to letter knowledge learning, Paris (2005) makes the useful distinction between learning constrained and unconstrained skills. Learning the letters of the alphabet is an example of a constrained skill as it is a small set of knowledge that can be mastered in a relatively short period of time. He contrasts this with vocabulary, which he suggested requires extensive and extended learning across a range of subjects.

In the South African context, Makaluza & Mpeta (2022) show using a randomised control trial that after one year, the Funda Wande literacy and maths teaching assistant (TA) intervention achieved a 0.44 standard deviation improvement in literacy and a 0.38 standard deviation in numeracy when compared to control schools. The main features of the Funda Wande TA intervention were: (1) the rigorous recruitment and selection of young people into the TA posts, (2) the provision of high-quality support material, (3) comprehensive training on the use of the material, and (4) continuous monitoring and support.

3. The intervention

Based on these findings, the Zazi iZandi project was developed. Working together with teams from Masinyusane, Funda Wande, The Learning Trust and the Binding Constraints Lab, the Zazi iZandi project was designed to measure the improvement in children's letter recognition after introducing teaching assistants into the classroom. Using baseline and endline results from the first subtest of the Grade 1 Early Grade Reading Assessment (EGRA), the findings could be used to build a case for using teaching assistants for instructional purposes.

The logic model in the program links the evidence-based design of the TA training program (1) with the learning outcomes (6) with the intermediary steps of fidelity of the training of TAs (2), monitoring and supporting TAs in the classrooms (3), the enactment of the games by the TAs (4), and the active participation of the children in the letter knowledge games (5) (Fig. 1).

The details of intervention program is described below. The intention of the six-to-nine-week intervention was to fast-track the learning of letter knowledge in isiXhosa and Sepedi. The letter sequence in the programme would reflect the frequency with which letters appear in the two languages.

The Guide details how the TAs were required to assess students, set up student groups, and initiate and manage the games and activities. The approach is built on WordWorks, Shine Literacy, Year Beyond and Molteno's Vula Bula.

The programme began with assessing students' letter knowledge using the following EGRA tool (Fig. 2):

Using a tracking system that the TAs were trained to use, the TAs were asked to indicate if the student got the letter correct or incorrect on the standard clipboard table provided. The purpose of the initial letter knowledge assessment was to establish a baseline of the learner's letter knowledge and to provide information to group students by initial letter knowledge. Thus, using the initial assessment results, TAs could group students with similar levels of letter knowledge.

Once the groups had been established, the TAs would meet each group separately for 20 min during their Home Language lessons. This was to happen on the carpet or rug at the back of the classroom in the Eastern Cape and sometimes in a separate room in Limpopo if space was available, with the children and the TAs sitting on the floor. Each small group's session was designed to cover three "new" letters as well as to repeat previous letters for consolidation. Each session consisted of the TAs managing a game or activity and recording progress on letter knowledge for each student on their assessment charts. Emphasis was placed in the *Guide* on making sure that the TAs worked at the right pace so that children were not left behind or made to feel overwhelmed.

TAs were required to mark attendance. They then began the session by signalling to the students the letters on the alphabet frieze cards that would be covered. The TA showed the letter on the cards and said the letter sound out loud. The students then said each letter individually and then as a group. The TA was required to ensure that each student said the letter name and sound correctly. The students often draw the letter either in the air or on paper. This was followed by the game or activity.

4. Implementation of the components

The following sections describe the steps in the implementation of the new TA programme.

4.1. Teacher/ school engagement - orientation

The initial stage of the implementation involves intervention managers visits to schools in early September 2023. The purpose of the visits to schools was to meet with Grade R and Grade 1 teachers as well as the school principal to brief them about the programme. The meeting was followed by an official circular from the provincial department of education. The briefing was followed by a question-and-answer session.

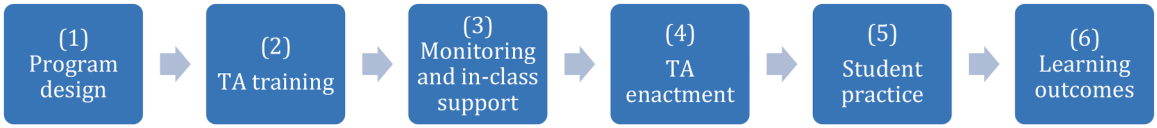


Fig. 1. Logic model of the teaching assistant training program intervention

l a w E Š f S I T e
d R m N t G P U š F
k j g r u O b A o A
E N i K J Y h M H r
p U W E T š y U s n
D Š p B y t g e M G

Fig. 2. EGRA letter knowledge subtask

Questions were raised by teachers about the role that teachers would play with the TAs, and about the potential space in classrooms for TAs to work with small groups. Although some teachers expressed concern about the fact that TAs would be watching teachers teach in the classroom, the overwhelming sentiment was that TAs would be getting involved directly in teaching and learning.

4.2. Training

The official training sessions for the TAs took place in both provinces on 7 and 9 September 2023. The training included an orientation to the programme manual and the use of the tracker tools, an orientation to the various games described in the manual, and the administering of the one EGRA question on letter knowledge. On the second day, the TAs were given the materials they would use for the letter knowledge games.

The following are some of the reflections on the training in Limpopo:

The training went well for the most part. The TAs were divided into four groups. When introducing a game or activity, the trainer would typically explain it then demonstrate with one or more of the TAs. They would then get a volunteer to role-play leading the game, in front of the entire group. The remaining TAs would then comment on how well the role play was done, what steps were missed and what could be improved upon. They would then go back into their respective smaller groups and practise the activity for 10–15 min. There was a reflection at the end of the training to check against the trainees' expectations at the start of the training. However, there wasn't much space to reflect on each activity before moving on to the next.

All but one of the TAs required to attend the training were present on the second day. There had been a mistake with one of the TAs on the first day as she was not eligible. The observers raised some concerns about the TA trainee's mastery of the workings of the assessment booklet and data collection booklets. One issue that surfaced during the training was the possibility that the TAs would be required to do tasks that they had formerly performed as education assistants (EAs) in the schools. At the end of the training, the trainers observed that the TA recruits lacked the confidence, energy and proactiveness that the NGO had come to expect in the previous project.

4.3. Materials

A very comprehensive guide was provided to all TAs on the project. In addition, each TA received the following resources at the end of the training session:

- ZZ Data Collection Booklet
- Letter Knowledge Assessment Booklet
- Blank training data collection and assessment sheets
- ZZ alphabet cards
- VulaBula letter/picture cards
- VulaBula mini alphabet chart
- Board game
- Big and small counters
- Dice
- Pencils, chalk and erasers
- Sand timer (not provided at training because of order delays)

4.4. Mentors

In the Eastern Cape interventions, the organisation provided five mentors for the 12 schools in the programme. They had organisation vehicles which made transport quite simple. Unfortunately, high incidences of robberies in the project areas meant they did not feel safe visiting schools with cell phones, so any data capturing they did was paper-based and would have to be captured electronically later. In Limpopo, one mentor was assigned full time to the project. She did not have a car and used public transport or arranged lifts from teachers from the various schools. On average, she managed to visit one school per day.

The mentors spent considerable time with TAs in their classrooms and interacted regularly with teachers and school leaders. In addition to providing a link between the school leadership and the project, the mentors played three distinct roles in the project. The mentors often modelled the practice, that is, they showed the TAs how to set up the games, demonstrated the initial assessment, assisted with the initial group selection and showed the TAs how to use the reporting booklets. The mentors played a critical role in ongoing monitoring of day-to-day sessions. Finally, the mentors played a critical role in maintaining the social and emotional level of commitment to the project. The mentor engagement was described as a "high touch presence".

4.5. Administrative data

One of the unusual components of the programme was the extensive collection of information on learning and the use of that information to guide instruction. In addition to internal project baseline and midline assessments with all students in the class to determine which students would be "graduated out" and identify the particular level of learning of each student, the programme required TAs to capture student information at each session. The focus of this information was on student letter mastery. This allowed the TA to record learning progress alongside teaching every set of games. Notwithstanding some confusion about whether correct new letter answers were to be colour coded or coded with a X, the data collection appeared, at least in the Eastern Cape appeared to be rigorous. Mentors would work closely with TAs to both track and analyse trends. Mentors themselves also had separate data collection forms that allowed them to report back regularly on growth in letter knowledge learning. The NGO head office analytics automatically checked all sources of data for errors and inconsistencies. After all data was cleaned the NGO management team, including the mentors, would do 'deep digging' to help pick up trends.

To gauge the level and quality of the TAs' enactment of the programme, we used a combination of secondary data analysis from the administrative data of the Eastern Cape and a field visit to the

programme in the intervention (Table A2).

The administrative data shows that, on average, Grade 1 students attended 11 letter name sessions with TAs and Grade R students around 13 sessions in the six-week period. Just under half of the Grade 1s "graduated" out of the programme, meaning that they reached the minimum threshold, while very few Grade R "graduated" by the end of the programme (Table 1).

Fig. 3 shows the uneven implementation of the intervention. The number of sessions per student was relatively low in the first and third week when the programme was getting started and began to tail off once the examinations began in the fourth quarter. Although the programme was designed to be implemented for the entire duration of Term 4, full implementation in fact occupied around six weeks.

5. Impact evaluation methodology

There is no standard framework for evaluating interventions. A widely used approach categorises impact research into five distinct phases (Medical Research Council, 2000). Phase 0 involves theory building and design conceptualisation. Phase 1 is associated with intervention modelling, where the design components are explored, and theories of change mapped. Phase 2 is referred to as an exploratory pilot trial phase in which constants and variables are described. This phase is best understood as feasibility trial. Phase 3 is normally associated with a randomised control trial with a scaled control group, allowing for causal inference. Phase 4, the final phase, involves studying population level implementation.

Within this framework, this study is most closely associated with a Phase 2 pilot in which the variables are described, the feasibility is assessed, and initial gains are measured. The pre- and post-testing provides an indication of a possible impact of the intervention, but given the nature of the sample, the findings cannot be generalised to the wider population. As the study and all associated data collected was initially not going to be used for publication, it was deemed unnecessary to distribute consent forms for students' parents to sign. However, students were required to ascent to participate in the external testing prior to commencement of the testing process.

5.1. Data sources and data collection

The[2] intervention sample for Zazi iZandi was drawn from two circuits of Primary School schools, one in the Eastern Cape and the other in Limpopo. At the Eastern Cape schools, the Language of Learning and Teaching (LoLT) is isiXhosa while Limpopo schools use Sepedi as their LoLT.

In a subsample of the intervention group, the external student assessment was conducted by an agency independent of the management of the project and the intervention agencies. The external evaluation agency assessed both Grade R and Grade 1 children using the Letter Recognition Subtask of the Grade 1 EGRA. The task involved students being timed on their ability to recognise 60 letter sounds in 60 s. The task was structured so that if students got the first five letter sounds incorrect, they were automatically timed out of the assessment. Demographic questions were asked at the start and end of the student assessment, including the child's language, gender and date of birth.

Table 1
Average number of letter knowledge sessions in the Eastern Cape

Grade	Graduated	Average No. of sessions	# of Students
Grade 1	Yes	9.19	507
Grade 1	No	12.81	615
Grade 1 Total		11.17	1122
Grade R	Yes	15.00	39
Grade R	No	12.72	737
Grade R Total		12.83	776

The external evaluation agency's baseline assessor training took place virtually 8 September 2023 with a refresher training for the endline held on 10 November 2023. Two experienced assessors were recruited for the baseline, one in Limpopo and one in the Eastern Cape. For the endline, the assessor from the Eastern Cape had to be replaced as they were committed to another study.

The external agency assessors spent one day at each school, assessing as many students as they could with specific instructions to complete a minimum of 12 assessments per grade for the baseline. Depending on how many students had been assessed during the baseline, this was then the target for the endline, ranging from 12 to 20 students per grade. While assessors largely arrived at the schools on time, there were times where they were delayed due to commitments like staff meetings, assemblies and principals wanting to be reintroduced to the project.

To select students during the baseline assessment, assessors completed a random selection form. They were instructed to line students up and assign them numbers, after which the form generates a list of random numbers which correspond to those which the assessor assigned to the students. Assessors then put numbered stickers on the selected students to make identification easier. Students were called into the assessment space one by one using their numbers to identify them. These numbers were used to generate the unique student IDs. For the endline assessments, assessors were given lists with previously assessed students.

The assessor who conducted the baseline assessments in the Eastern Cape was not available during the endline and while Data Collection Agency data has an extensive roster of fieldworkers, finding those in the Eastern Cape with drivers' licences was a challenge. Although the fieldworker that was initially recruited for the endline had worked on a similar project before, it was evident from the on-the-ground staff that this assessor was not following the research protocols and that this was ultimately interfering with student outcomes. After withdrawing the assessor from the field, they were replaced by a highly experienced assessor who was comfortable working alone.

According to the fieldwork report, the main constraint that the data collection team faced was the length and structure of the Primary School Day. Across the schools, Grade R and 1 students largely left school after 13h00. Additionally, Primary School schools finish early on Fridays, limiting the time even further.

5.2. Ethics clearance

The study received ethics clearance from the University of the Witwatersrand Non-Medical Ethics Committee.

6. Finding

For the following section, we focus on investigating the impact of the intervention on letter sound knowledge measured by the change in correctly identifying letters from baseline to endline.

6.1. Student performance data

As we outlined above, the independent external Data Collection Agency sample of schools contains 10 schools from the Eastern Cape and Limpopo. Table 2 shows the breakdown of the number of students that were tested by the independent Data Collection Agency at baseline and endline. A total of 289 Grade R and Grade 1 students were tested for the baseline. However, only 266 students were re-tested in the endline. As is illustrated in the third column in Table 2, attrition was spread across the schools.

To establish the gains experienced by the students, we need to be

[2] This section is drawn from the external agency's fieldwork report.

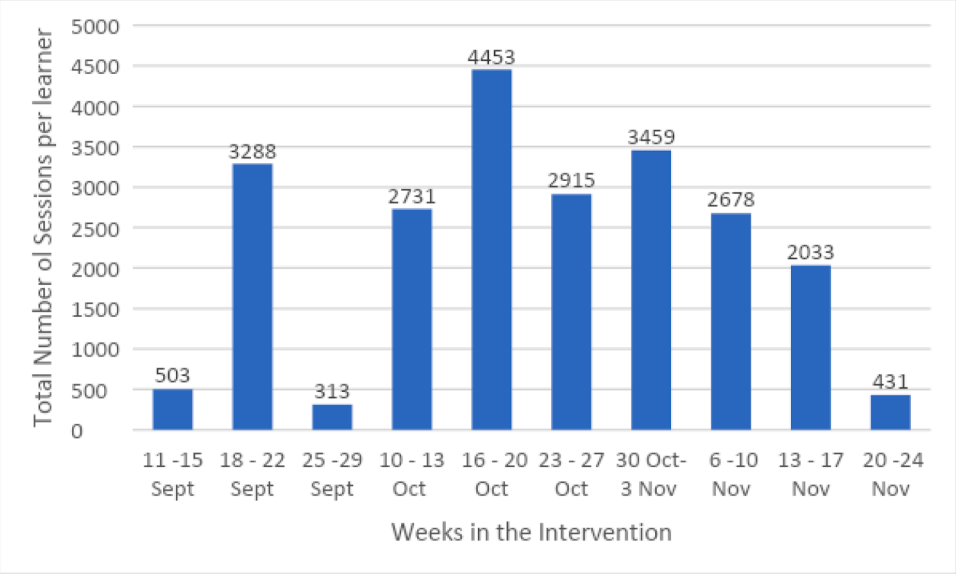


Fig. 3. Number of sessions per student by week

Table 2
Data Collection Agency student data (number of students tested)

School	Baseline	Endline missing data
1 Primary School	30	8
2 Primary School	30	1
3 Primary School	30	1
4 Primary School	35	1
5 Primary School	27	3
6 Primary School	28	5
7 Primary School	26	1
8 Primary School	28	1
9 Primary School	31	
10 Primary School	24	2
Total	289	23

able to follow the same sample of students over the study period. Therefore, we exclude 46 observations (the same 23 students from baseline and endline for whom we do not have endline data) from the total sample of 578 student observations, which leaves us with a complete sample of baseline and endline student performance data of 532. This way, any changes in the average number of correct letters reflects changes experienced by the same sample of students and is not a function of a different composition of students in the baseline and endline samples.

Additionally, we exclude 48 student observations of one particular Primary School (12 Grade R and 12 Grade 1 students in baseline and endline) from the sample as we suspect that one specific assessor inflated the performance of the students in the endline survey (see Table A1 in the appendix for justification). Therefore, we continue our analysis with 484 observations of the independent Data Collection Agency dataset of student letter knowledge for whom we have a complete set of baseline and endline data.

We test if the missing and excluded learner observations are correlated with baseline characteristics and baseline learner performance. Only the endline performance data of the 23 learners are missing completely, but none of the baseline performance nor any of the individual learner characteristics (see Table A3 in the appendix). As we report in Tables A4–A6 (appendix), the excluded learners (including the learners from the excluded Primary School 5) are on average older, more likely to be in Grade 1, and have attempted more letters and got more letters correct at baseline. However, only the difference of a higher number of correct letters is statistically significant for excluded Grade R learners but is otherwise not correlated with any of the individual learner characteristics. More importantly, the majority of excluded learners attended schools in the Eastern Cape province.

6.2. *Data Collection Agency and administrative data*

To complement the independently collected student performance data from the Data Collection Agency, the intervention organisations provided administrative data collected by the TAs. This group of schools was significantly larger and included a larger number of students per school. As such, the administrative data provided additional insights into the mechanisms of the intervention. TAs collected performance data at baseline for a total of 1 894 Grade R and Grade 1 students. Unfortunately, endline data was available for only 1 059 students. The main reason for this was that many of the students had already stopped coming to school in the EC in the week scheduled for TA-administered endline there, which was the week following the independent data collection. As the modus of the testing was different to the one-minute testing done by the Data Collection Agency using the standard EGRA tool, the administrative data is not easily comparable to the Data Collection Agency data.

Tables 3 and 4 show a comparison of student performance data as collected by the independent Data Collection Agency and the TAs by province and grade.

Both data sets show that students on average experienced gains over the study period. This is assuring, as it confirms that gains were observed through two different and independent data collection processes. The baseline data especially reflects comparable starting points across provinces and grades. Thus, data collected by both the independent Data Collection Agency and the TAs are comparable before the intervention was implemented. However, the magnitude of gains experienced by the students differs between the two data collection processes. The data collected by the TAs shows significantly larger gains between the baseline, midline and endline.

The midline data collection of the TAs was closer in time to the endline data collection of the independent Data Collection Agency. Irrespective, the gains captured by the TAs between baseline and midline still exceed the gains captured by the independent Data Collection Agency by close to double in magnitude.

This is partly due to different time limits that were imposed by the TAs and the Data Collection Agency. Specifically, the TAs allowed students some additional time to complete the tasks compared to the 1-

Table 3
Average number of correct letters: Data Collection Agency vs administrative data (Limpopo Province)

	Data collection Agency			Administrative data from TAs				
	Base Line	Endline	Change BL-EL	Base Line	Mid Line	End Line	Change BL-ML	Change BL-EL
Grade R	2.7	6.3	3.7	4.8	11.6	18.9	6.8	14.1
Grade 1	24.8	30.8	6.0	22.5	31.3	41.2	8.8	18.7

Table 4
Average number of correct letters: Data Collection Agency vs administrative data (Eastern Cape Province)

	Data collection Agency			Administrative data from TAs				
	Base line	End line	Change BL-EL	Base line	Mid line	End line	Change BL-ML	Change BL-EL
Grade R	5.2	11.7	6.5	7.3	17.6	26.6	10.3	19.3
Grade 1	26.1	33.2	7.1	26.7	39.8	47.1	13.1	20.4

minute time limit imposed by the Data Collection Agency. Additionally, the TA data includes a larger number of students and schools compared to the subset of students and schools that were tested by the Data Collection Agency.

To test the comparability of the two datasets, we restrict the performance data to the same sample of students that were tested by the Data Collection Agency as well as by the TAs in the Eastern Cape Province. Specifically, using Spearman ranking coefficient, we investigate the correlation between the students' performances at baseline and for the same sample of students at midline and endline.

Table 5 above reports a strong correlation between the Data Collection Agency baseline data and the TAs' administrative baseline data with a correlation coefficient of 0.86. Thus, even with slight differences in the magnitude of the measured performances, the two datasets move in the same direction. Unfortunately, this strong correlation between the two data sets diminishes in the midline and endline. The TAs' midline and endline data that allows students extra time to identify letters is only moderately correlated with the endline data of the independent data collection process of Data Collection Agency (0.57 and 0.62 respectively).

Therefore, to provide policy relevant findings that allow for an analysis of the intervention on students' performance, we proceed with the analysis of the data collected by the independent external Data Collection Agency. Additionally, as the independent Data Collection Agency implemented the 1-minute time limit, the findings can be assessed against the DBE benchmarks. Nevertheless, the TA data provides empirical support that the Data Collection Agency data is accurate in measuring student performance at baseline with the assumption that the independence of the Data Collection Agency enumerators provides a more accurate picture of the performance at the endline as well.

6.3. Main finding

Table 6 reports the average gains that students made both in Grade R and in Grade 1 in the intervention and their respective standard deviations (in brackets). Students in both grades managed to correctly identify around 5–7 letters more after the intervention compared to their baseline performances. In the Eastern Cape, Grade 1 students improved by seven letters which represents .36 of a Standard Deviation. In terms of Standard Deviations, the gains are of similar magnitude in Limpopo (.32 of a SD), save that the students were slightly lower at baseline. For the Grade Rs, the gains are much larger across both provinces (.67 of a SD) as they more than doubled the number of correct letters. However, one needs to keep in mind that they come off a much lower baseline for both the Eastern Cape and Limpopo compared to the Grade 1 learners.

Table 5
Spearman rank correlation of data collection agency and TAs performance data for the same students of Eastern Cape schools

	Administrative data (TAs)	Administrative data (TAs)	Administrative data (TAs)
	Baseline	Midline	Endline
Data Collection Agency (baseline)	0.86		
Data Collection Agency (endline)		0.57	0.63

According to the DBE website for the Sesotho-Setswana language group, the benchmark indicates that by the end of Grade 1, all students should be able to identify 40 correct letter sounds per minute. By the end of Grade 2, all students should be able to read at least 40 correct words per minute (DBE, no date). Similar benchmarks for Grade 1s in Nguni languages have been published with the same Grade 1 40 letter benchmark.

Table 7 reveals that the intervention had positive effects both for students that had no prior knowledge of letters, and significantly increased the proportion of students that meet the Department of Basic Education benchmark letter knowledge of 40 letters per minute. The increase of the proportion of students that meet the benchmark for the end of Grade 1 from 29 % to 42 % is educationally meaningful. Equally important is the reduction of students that could not identify a single letter at baseline. However, the rather low % change seems to indicate that non-identification of letters is potentially not a pedagogic problem but could rather result from severe learning challenges.

While averages and learning thresholds are useful measurements of overall achievements, they are likely to mask important policy relevant nuances of the intervention along the performance distribution. We therefore unpack the changes along the performance distribution.

We start our investigation by looking at the number of letters that the student attempted during the 1-minute test. The dashed line reflects the number of attempted letters at baseline while the solid line shows the number of attempted letters at endline. It is clear that a large proportion of the Grade R students attempted at least some letters after the intervention compared to the majority attempting only a few letters at baseline. Additionally, we see an increase of attempted letters across the distribution with some students attempting all letters by the time of the endline. For Grade 1 students, we see a more proportional improvement of the number of letters attempted (Fig. 4).

A very similar pattern is observed for the number of correctly identified letters. Significantly fewer Grade R students cannot identify a single letter correctly while an increasing portion of the same sample manage to identify 5 letters and more correctly.

Teaching at the right level approach is reflected in the parallel shift of the two cumulative curves shown in Fig. 5 above. This could reflect the teaching approach of grouping students by initial letter knowledge and therefore teaching at the right level. Thus, this intervention seems to allow students across the performance level to benefit equally from the

Table 6
Main findings: average number of letters correct at baseline and endline for Grade R and Grade 1 (Data Collection Agency)

Sample	Number of learners	Baseline average number of correct letters (SD)	Endline average number of correct letters (SD)	Difference in number of correct letters	Cohen's d
Full Sample					
Grade R & 1	242	14.1 (17.79)	20.3 (18.77)	6.2	.30
Grade 1 (only)	122	25.3 (19.24)	31.7 (18.95)	6.4	.34
Grade R (only)	120	3.8 (5.2)	8.7 (8.9)	4.9	.67
Eastern Cape province / isiXhosa					
Grade R & 1	96	14.8 (18.2)	21.6 (17.6)	6.8	.38
Grade 1 (only)	44	26.1 (20.8)	33.2 (18.4)	7.1	.36
Grade R (only)	52	5.2 (6.8)	11.7 (8.8)	6.5	.83
Limpopo province / Sepedi					
Grade R & 1	146	14.7 (17.6)	19.4 (19.5)	4.7	.25
Grade 1 (only)	78	24.8 (18.4)	30.8 (19.3)	6.0	.32
Grade R (only)	68	2.7 (2.96)	6.3 (8.2)	3.6	.58

Table 7
Percentage of students with zero or at or above 40 on baseline and endline for Grade 1

	Baseline			Endline		
	#	Total	%	#	Total	%
Zero correct letters	11	122	9.0 %	8	122	6.6 %
1 – 39 correct letters	76	122	62.3 %	63	122	51.6 %
40 and more correct letters	35	122	28.7 %	51	122	41.8 %

intervention.

As our Fig. 6 shows, not only have students attempted a higher number of letters, but they also managed to get a larger share of their attempted letters correct. Clearly, the higher number of correct letters could have simply been that they attempted more and had lucky guesses. However, as Fig. 6 illustrates for both, Grade R and Grade 1 students, the share of correct letters of attempted letters increased from baseline to endline. Particularly the Grade 1 students experienced a significant increase in getting all or at least a larger share of their attempted letters correct.

6.4. Regression analysis

The above analysis has shown that the average changes for the same sample of students was experienced across the distribution. In the following, we investigate if the data follows similar educational outcomes as have been identified in the South African education literature before. This will support our confidence in the validity of the test data and might suggest potential areas that need to be investigated in future studies that assess the impact of this intervention and its heterogeneous treatment effects.

In that context we test if the positive change in attempted letters, correct letters and the share of correct letters of attempted letters is correlated with individual student characteristics (gender and age of

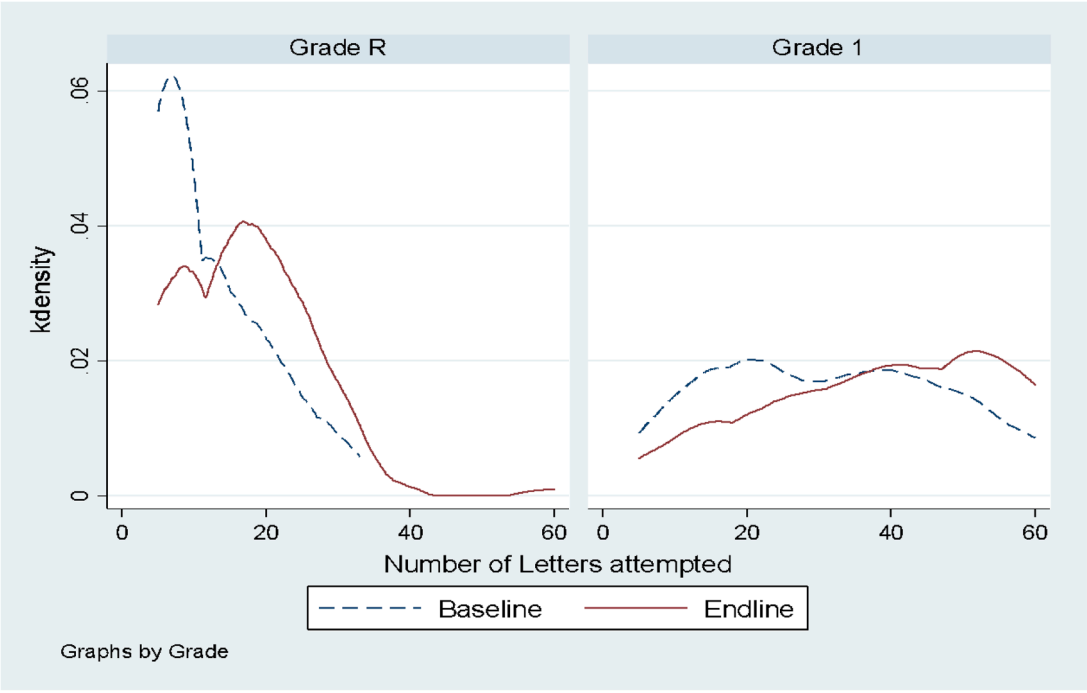


Fig. 4. Number of letters attempted baseline vs endline (by grade)

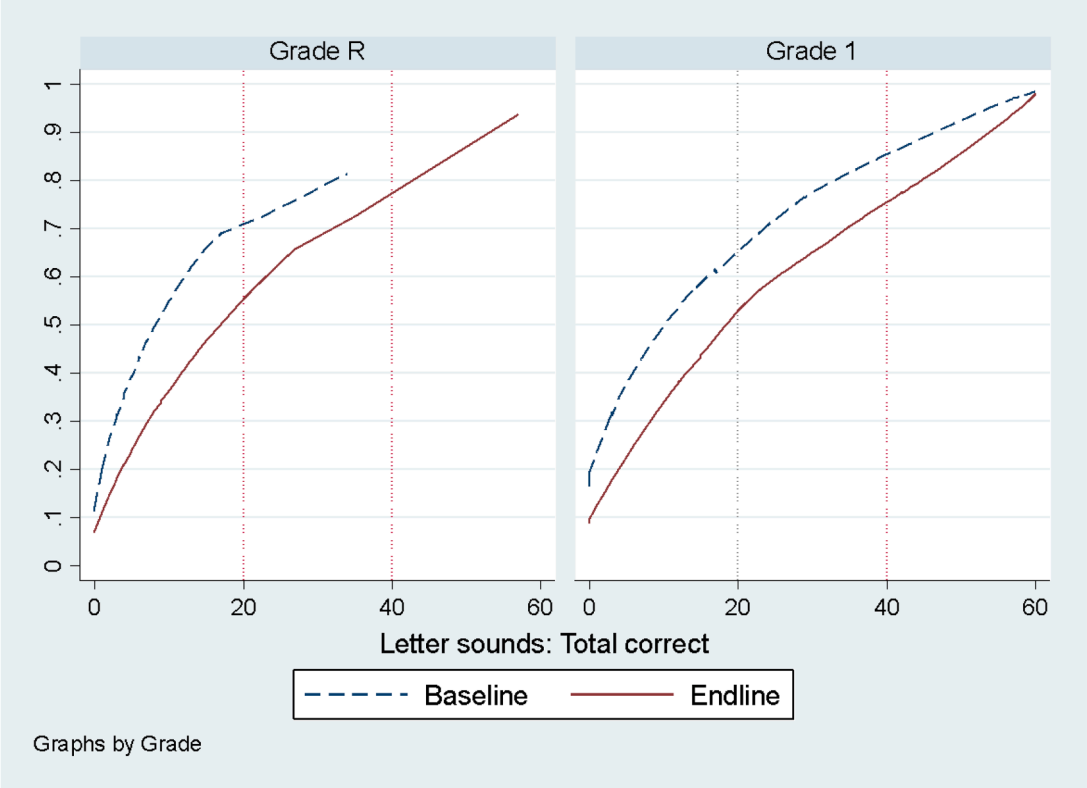


Fig. 5. Cumulative distribution of the number of letters correctly identified baseline vs endline (by Grade)

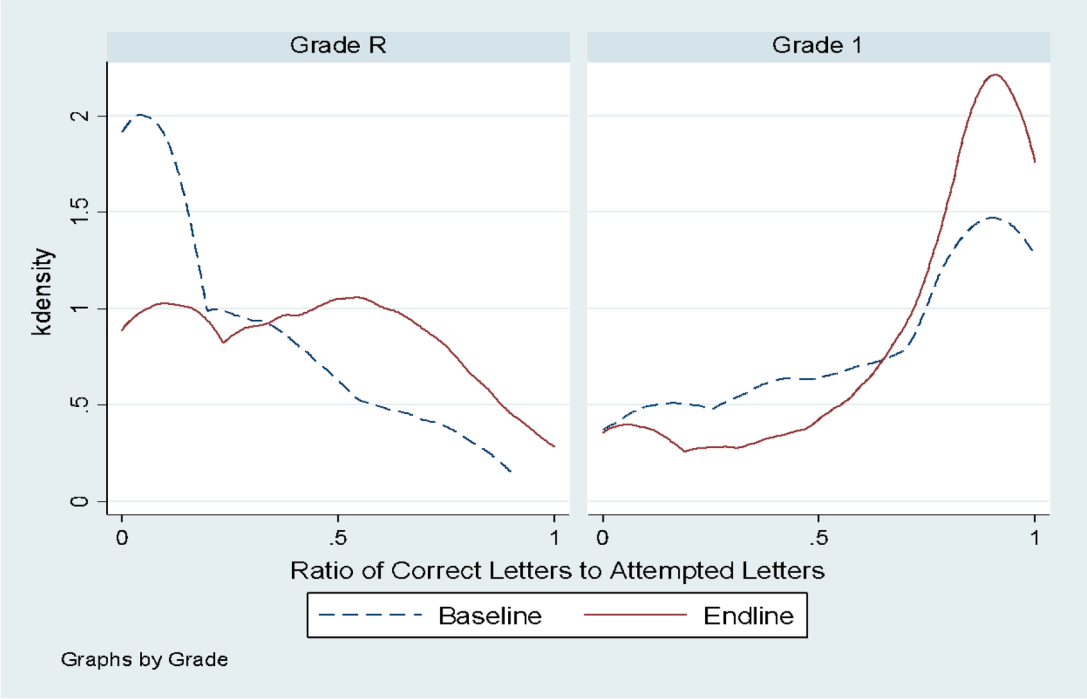


Fig. 6. Share of correct letters to attempted letters baseline vs endline (by Grade)

learner), household characteristics (number of books in the household) and therefore account for the gains experienced by the students. To account for the fact that students are clustered within schools, we include cluster adjusted robust standard errors by clustering at the school level. The coefficient are reported as Z-scores using the pooled standard deviations of the full sample (including Grade R and Grade 1 learners).

Irrespective of the inclusion of additional controls, the effect of the intervention remains robust as the point estimate of the *Endline* variable does not change significantly and remains highly statistically significant

for attempted letters and correct letters. The fact that the interaction term *Endline*Gr 1* is statistically not significant implies that Grade R (Gr R) and Grade 1 (Gr 1) students experienced similar absolute gains between the baseline and the endline.

Female students as well as students living in a household with books were more likely to attempt letters and identify letters correctly.

The final regression (columns 5 and 6 in Table 8) show the ratio of correct letters to attempted letters. Not only did students attempt more letters in the endline test, but they also managed to get more correct. All in all, the intervention improved the success rate in letters attempted correctly in both grades. Nevertheless, the intervention was most successful for Grade R students in improving the correct identification of attempted letters.

It is noteworthy that female students not only attempted more letters compared to their male peers, but also have a higher share of getting attempted letters correct.

6.5. Heterogeneous treatment effects of the teacher assistant pilot study

In line with findings of the education literature in South Africa (Cilliers et al 2020; Moholwane, et al 2023; also see a review of research in Basharati et al 2021) we unpack the potential effect of the treatment further by investigating possible heterogenous effects. Specifically, given the significance of the student's gender (Zuze & Reddy, 2014) and their household characteristics (Heaton et al, 2014) on learning outcomes, we investigate if the gender and having a household environment that is conducive to learning letters (measures by the number of books at home) affected changes in Letter Knowledge. While the regression includes the full set of control variables as is used in Model 2 above, we only report on the interaction effects in Table 9.

As is shown in Table 9, neither the gender of the student nor the number of books in the household are correlated with the gains experienced by the students over the study period.

7. Discussion

The pilot study was designed to assess the potential feasibility and potential effectiveness of a teaching assistant programme focused on teaching letter knowledge. Judged against several criteria, the pilot model and programme are promising.

First, the Zazi iZandi programme design is consistent with current educational change thinking. Specifically, the programme makes use of three approaches: 1) the use of a structured pedagogy, 2) teaching at the right level, and 3) learning through play. Structured pedagogy (UNICEF, 2020), while not without its critics, is an approach that focuses on emphasising instructional designed learning packages. These packages incorporate aligned and coherent learning materials, training, assessment and ongoing support with carefully planned scope and sequence of content to be learnt. The Zazi programme certainly incorporates these design features. The idea of teaching at the right level (Banerji & Chavan, 2016) is not new in education, and it is one of the central principles of Vygotsky's zone of proximal development. The idea is that learning is optimised when the student is provided a task in a zone that is neither too easy nor too complex, but at the right level, and can be mastered with appropriately targeted instruction. The initial assessment of letter knowledge, and the grouping of students based on the performance on assessment of their letter knowledge and then the tailoring of instruction to the average level of the small group is consistent with the principle of teaching at the right level. The principle of learning through play is also not new (UNICEF, 2018), with its origins in the work of Maria Montessori and John Dewey but provides a strong evidence-based principle to enhance learning outcomes for young children. The research shows that play enhances levels of engagement, provides a healthy social and emotional climate for learning, and ensures a balance between kin-aesthetic and cognitive development. The set of eight games that are at the core of the Zazi iZandi programme embodies the learning through play idea.

Second, the programme addresses important implementation criteria. In terms of the Technology Adoption Model (TAM) (Marangunić & Granić 2015) on the take-up of new technologies, intervention technologies are most likely to be successfully institutionalised if two conditions are met, that is, that the new technology is easy-to-use/easy to learn-to-use and that the users perceive them as effective in their daily work. The evidence from the Eastern Cape version of the pilot suggests that these have been met.

Third, the programme addresses important implementation criteria. It focuses on an early learning gap that has been identified in South African research as a crucial obstacle to literacy learning outcomes, i.e. very slow progress in letter knowledge development (Ardington et al, 2021). Letter knowledge has been shown to be both associated with and

Table 8
Regression output of the number of letters attempted and correct letters (Z-scores)

Variables	(1) Attempted letters	(2) Attempted letters	(3) Correct letters	(4) Correct letters	(5) Correct/ attempted letters	(6) Correct/ attempted letters
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Grade R (Endline)	0.263***	0.262***	0.254**	0.253**	0.501**	0.500***
	(0.0855)	(0.0865)	(0.0889)	(0.0874)	(0.185)	(0.171)
Grade 1 (Baseline)	1.111***	1.010***	1.114***	0.982***	1.194***	1.063***
	(0.141)	(0.133)	(0.137)	(0.126)	(0.101)	(0.114)
Endline*Grade 1	0.0997	0.0939	0.0791	0.0744	–0.297*	–0.297
	(0.205)	(0.207)	(0.192)	(0.194)	(0.168)	(0.171)
Female		0.399***		0.356***		0.298***
		(0.0777)		(0.0680)		(0.0679)
Student age		0.0961		0.126		0.140*
		(0.0716)		(0.0762)		(0.0727)
Books at home (yes=1)		0.205**		0.170**		0.0419
		(0.0714)		(0.0671)		(0.0773)
Limpopo province		–0.0854		–0.127		–0.244
		(0.108)		(0.128)		(0.163)
Constant	–0.776***	–1.547***	–0.772***	–1.623***	–0.821***	–1.578***
	(0.0520)	(0.362)	(0.0461)	(0.377)	(0.112)	(0.371)
Observations	484	484	484	484	484	484
R-squared	0.393	0.460	0.387	0.445	0.313	0.360

Cluster adjusted (school level), robust standard errors in parentheses
*** $p < 0.01$
** $p < 0.05$
* $p < 0.1$

Table 9
Regression output of heterogeneous effects (Z-scores)

	(1) Att letters	(2) Att letters	(3) Att letters	(4) Corr letters	(5) Corr letters	(6) Corr letters	(7) Corr/ att letters	(8) Corr/ att letters	(9) Corr/ att letters
	Full	Gr R	Gr 1	Full	Gr R	Gr 1	Full	Gr R	Gr 1
Endline* Female	–0.0115	0.0293	–0.0319	0.0333	0.0865	–0.000345	0.0356	0.0815	–0.0476
	(0.152)	(0.124)	(0.241)	(0.127)	(0.0732)	(0.251)	(0.127)	(0.168)	(0.244)
Endline* books at home	0.137	0.143	0.114	0.130	0.108	0.149	0.125	0.194	0.143
	(0.134)	(0.159)	(0.286)	(0.132)	(0.131)	(0.297)	(0.150)	(0.276)	(0.260)
Obs	484	239	245	484	239	245	484	239	245

Cluster adjusted, robust standard errors in parentheses,
*** $p < 0.01$, ** $p < 0.05$,
* $p < 0.1$

more recently causally linked to building an understanding of the links between phonemes and graphemes, word decoding and ultimately reading for comprehension (Noel Foulin, 2005). That said, letter knowledge is a necessary but far from sufficient condition for reading progress. Without all the other building blocks, letter knowledge in-and-of-itself is unlikely to address the reading crisis faced by the country.

Fourth, although the teacher assistants have neither the level of training nor the experience of most fully qualified teachers, the narrow focus on improving students' letter knowledge may enable young people to make a meaningful educational contribution. The shift of the work of education assistants in the first years of the national rollout from administrative tasks in schools to educational tasks ensures that two of the key government objectives - extending work opportunities and improving reading outcomes - could be addressed concurrently.

8. Conclusion

The findings of the Phase 2 pilot study of the Zazi Zandi intervention, notwithstanding the challenges encountered, provide a strong basis for moving to the next evaluation phase. Given the findings of the pilot study, the feasibility associated with the implementation of the programme, and the size of the learning gains measured between pre-test and post-test by the external evaluator, there is a strong case for moving toward a full impact evaluation, that is a large-scale cluster randomised control trial. Such a Phase 3 trial would provide policy relevant and generalisable evidence of the cost-effectiveness of the programme model using a representative sample of schools.

While the gains reported between pre-test and post-test are statistically significant and educationally meaningful, we cannot rule out the possibility that the changes could be due to the maturation of students or typical learning gains at the end of the year. This is unlikely however, given the short intervention period. Nevertheless, to rule out this possibility it is necessary to incorporate a counterfactual - a control group of schools that will not be exposed to the intervention - in the next phase of the research. A counterfactual group would also allow for the investigation of heterogeneous treatment effects, moderators and mitigators of the intervention on student outcomes.

It is suggested that the Phase 3 trial should test two versions of the programme model, one with the current six-nine-week intervention and a second version with twice duration, i.e., 12–18 weeks or two school terms. If possible, the trial could take place in 2024, starting in Term 2 with testing the intervention model with the longer dosage, followed in Term 3 with the commencement of the shorter dosage model. The baseline testing for the control and two intervention programme groups would occur at the beginning of Term 2 with the endline completed by the end of Term 3. As with this pilot study, there should be a separation between the implementing agents and the agents tasked with conducting the baseline and endline testing.

CRediT authorship contribution statement

Brahm Fleisch: Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition. **Volker Schöer:** Writing – original draft, Methodology, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Funding for this study was made available by Allan and Gill Gray Philanthropies Africa and DG Murray Trust. The authors are solely responsible for the contents, however.

Appendix

Table A1
School level average number of letters correct baseline and endline for Grade R and Grade 1

	Baseline	#	Endline	#	Change
Grade 1	25.34	150	32.30	134	6.96
1 Primary School	23.53	15	19.13	8	–4.41
2 Primary School	22.00	15	31.33	15	9.33
3 Primary School	22.73	15	33.43	14	10.70
4 Primary School	16.76	21	24.20	20	7.44
5 Primary School	**18.00**	**15**	**33.83**	**12**	**15.83**
6 Primary School	37.75	12	41.60	10	3.85
7 Primary School	40.21	14	45.36	14	5.14

(*continued on next page*)

Table A1 (*continued*)

	Baseline	#	Endline	#	Change
8 Primary School	17.50	14	21.92	13	4.42
9 Primary School	29.24	17	34.65	17	5.41
10 Primary School	33.08	12	38.36	11	5.28
Grade R	4.19	139	12.49	132	8.30
1 Primary School	2.73	15	6.57	14	3.84
2 Primary School	5.87	15	14.29	14	8.42
3 Primary School	2.13	15	5.67	15	3.53
4 Primary School	2.29	14	4.64	14	2.36
5 Primary School	**2.83**	**12**	**55.33**	**12**	**52.50**
6 Primary School	11.75	16	16.15	13	4.40
7 Primary School	4.00	12	6.50	12	2.50
8 Primary School	2.29	14	8.69	13	6.41
9 Primary School	2.79	14	2.50	14	–0.29
10 Primary School	4.08	12	9.73	11	5.64
Grand Total	15.17	289	22.47	266	7.30

5 Primary School seems to be an outlier in the entire group of schools tested by the independent Data Collection Agency with the highest gains compared to all other schools and grades. The gains of 55 additional letters in the endline of Grade R students is particularly worrying. *5 Primary School* was also the only school tested by one particular enumerator. All other schools were tested by two of the other enumerators. For those two enumerators, the gains are relatively comparable across schools and grades. The inflated gains in *5 Primary School* therefore are more likely to be due to enumerator effects rather than intervention effects.

Table A2
Zazi iZandi pilot descriptive statistics

	Peri-urban Eastern Cape	Rural limpopo
Implementing partner	Masinyusane	Funda Wande
Language	isiXhosa	Sepedi
Number of schools	12	13
Number of TAs / Classes	51	26
Number of mentors	5	1
Number of students	1,901	739
Average class size	45	25
Median class size	38	22
Number of student sessions	11,334	19,209
Average dosage: sessions per student	12	26
Median dosage: sessions per student	11	23

Table A3
Missing observations

Variable	Missing (baseline)	Missing (endline)	Total	Percent Missing
Learner Characteristics				
Id learner	0	0	578	0.00
Learner Sex	0	0	578	0.00
Learner age	0	0	578	0.00
Grade	0	0	578	0.00
Books at home	0	0	578	0.00
Province	0	0	578	0.00
Outcome variables				
Number of attempted letters	0	23	578	3.98
Number of correctly identified letters	0	23	578	3.98

Table A4
Zazi iZandi pilot descriptive statistics of study sample (baseline)

Variable	All observations ($N = 289$)		Study sample ($N = 242$)		Excluded learners ($N = 47$)	
	Mean	SD	Mean	SD	Mean	SD
Female	0.54		0.54		0.53	
Age	5.62	0.81	5.59	0.81	5.77	0.76
Books at Home (Yes = 1)	0.61		0.60		0.66	
Grade 1	0.52		0.51		0.57	
Limpopo	0.52		0.60		0.09	
Number of attempted letters (baseline)	22.58	16.44	22.21	16.25	24.45	17.44
Number of correctly identified letters (baseline)	15.17	17.99	14.70	17.79	17.57	19.03

Table A5
Probit regression of missing endline outcome variables (baseline performance)

Variables	(1) Full sample	(2) Full sample	(3) Full sample	(4) Grade R	(5) Grade R	(6) Grade R	(7) Grade 1	(8) Grade 1	(9) Grade 1
Attempted letters (BL)	0.00446		–0.00692	0.0132		–0.0324	0.000629		0.000562
	(0.00527)		(0.0189)	(0.0154)		(0.0309)	(0.00716)		(0.0274)
Correct letters (BL)		0.00475	0.0108		0.0359*	0.0691*		0.000525	5.97e–05
		(0.00478)	(0.0172)		(0.0197)	(0.0375)		(0.00614)	(0.0235)
Constant	–1.087***	–1.059***	–0.996***	–1.233***	–1.236***	–0.986***	–0.936***	–0.929***	–0.935***
	(0.152)	(0.118)	(0.209)	(0.241)	(0.166)	(0.283)	(0.259)	(0.197)	(0.360)
Observations	289	289	289	139	139	139	150	150	150

Standard errors in parentheses,
*** $p < 0.01$, ** $p < 0.05$,
* $p < 0.1$

Table A6
Probit regression of missing endline outcome variables (baseline characteristics)

	(1) Full Sample	(2) Full Sample	(3) Full Sample	(4) Full Sample	(5) Full Sample	(6) Full Sample	(7) Full Sample	(8) Full Sample	(9) Full Sample	(10) Grade R only	(11) Grade 1 only
Female	–0.0117	–0.0178	–0.0297	–0.0373	–0.0673	–0.0844	–0.0835	–0.0835	–0.0835	–0.180	–0.0105
	(0.177)	(0.178)	(0.179)	(0.180)	(0.199)	(0.210)	(0.210)	(0.210)	(0.210)	(0.296)	(0.303)
Learner age		0.163	0.153	0.220	0.0179	0.0130	0.0137	0.0137	0.0137	0.0245	–0.0420
		(0.115)	(0.116)	(0.184)	(0.173)	(0.174)	(0.175)	(0.175)	(0.175)	(0.264)	(0.250)
Books at home			0.112	0.120	–0.0311	–0.0303	–0.0298	–0.0298	–0.0298	–0.0781	0.0476
			(0.187)	(0.187)	(0.213)	(0.213)	(0.213)	(0.213)	(0.213)	(0.291)	(0.324)
Grade 1				–0.135	0.205	0.173	0.174	0.174	0.174		
				(0.285)	(0.304)	(0.327)	(0.327)	(0.327)	(0.327)		
Limpopo					–1.454***	–1.453***	–1.457***	–1.457***	–1.457***	–1.159***	–1.626***
					(0.247)	(0.247)	(0.248)	(0.248)	(0.248)	(0.360)	(0.348)
Attempted letters (BL)						0.00199	0.00508	0.00508	0.00508	–0.0102	0.0124
						(0.00751)	(0.0213)	(0.0213)	(0.0213)	(0.0323)	(0.0311)
Correct letters (BL)							–0.00298	–0.00298	–0.00298	0.0299	–0.0115
Female							(0.0192)	(0.0192)	(0.0192)	(0.0394)	(0.0268)
	–0.977***	–1.896***	–1.904***	–2.212**	–0.648	–0.641	–0.669	–0.669	–0.669	–0.682	–0.220
	(0.129)	(0.666)	(0.664)	(0.939)	(0.898)	(0.898)	(0.918)	(0.918)	(0.918)	(1.376)	(1.630)
Observations	289	289	289	289	289	289	289	289	289	139	150

Standard errors in parentheses,
*** $p < 0.01$,
** $p < 0.05$, * $p < 0.1$

References

Ardington, C., Wills, G., Pretorius, E., Mohohlwane, N., & Menendez, A. (2021). Benchmarking oral reading fluency in the early grades in Nguni languages. *International Journal of Educational Development, 84*, Article 102433.

Banerji, R, & Chavan, M (2016). Improving literacy and math instruction at scale in India's primary schools. *Journal of Educational Change, 17*(4), 453–475.

Besharati, N. A., Fleisch, B., & Tsotsotso, K. (2021). Interventions to improve student achievement in South Africa: A systematic meta-analysis. *Systematic Reviews of Research in Basic Education in South Africa. African Sun Media*, 27–67.

Cilliers, J., Fleisch, B., Prinsloo, C., & Taylor, S. (2020). How to improve teaching practice?: An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources, 55*(3), 926–962.

Heaton, T. B., Amoateng, A. Y., & Dufur, M. (2014). Race differences in educational attainment of youth aged 7–18 in post-apartheid South Africa: The role of family structure, resources and school quality. *South African Review of Sociology, 45*(1), 101–121.

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological science, 23*(6), 572–577.

Makaluza, N., & Mpeta, B. (2022). Two birds with one stone: Improving Grade 1 learning outcomes using Funda Wande teaching assistants in Limpopo 2021-2022. *Interventions*, 146.

Marangunić, N., & Granić, A. (2015). Technology acceptance model: A literature review from 1986 to 2013. *Universal Access in the Information Society, 14*, 81–95.

McBride-Chang, C. (1999). The ABCs of the ABCs: The development of letter-name and letter-sound knowledge. *Merrill-Palmer Quarterly (1982-)*, 285–308.

Medical Research Council. (2000). *A framework for the development and evaluation of RCTs for complex interventions to improve health*. London: Medical Research Council. MetaMetrics.

Mohohlwane, N., Taylor, S., Cilliers, J., & Fleisch, B. (2023). Reading skills transfer best from home language to a second language: Policy lessons from two field experiments in South Africa. *Journal of Research on Educational Effectiveness*, 1–24.

Ngema, M. (2021). Exploring teachers' challenges in teaching reading proficiency to isiZulu Foundation Phase students at rural schools in KwaZulu-Natal. *Journal for Language Teaching, 55*(2), 195–217.

Noel Foulin, J. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing, 18*, 129–155.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*(2), 184–202.

Spaull, N., & Pretorius, E. (2019). Still falling at the first hurdle: Examining early grade reading in South Africa. *South African Schooling: The Enigma of Inequality: A Study of the Present Situation and Future Possibilities* (pp. 147–168).

Sunde, K., Furnes, B., & Lundetræ, K. (2020). Does introducing the letters faster boost the development of children's letter knowledge, word reading and spelling in the first year of school? *Scientific Studies of Reading, 24*(2), 141–158.

Treiman, R. (2013). Knowledge about letters as a foundation for reading and spelling. *Handbook of Orthography and Literacy* (pp. 581–599).

UNICEF. (2018). *Learning through play*. https://www.unicef.org/sites/default/files/2018-12/UNICEF-Lego-Foundation-Learning-through-Play.pdf.

UNICEF. (2020). *Structured Pedagogy: For real-time equitable improvements in learning outcomes Working paper*. https://www.unicef.org/esa/documents/structured-pedagogy.

Zuze, T. L., & Reddy, V. (2014). School resources and the gender reading literacy gap in South African schools. *International Journal of Educational Development, 36*, 100–107.