

Funda Wande Eastern Cape Evaluation

ENDLINE REPORT

APRIL 2022



Contents

INTRODUCTION	1
INTERVENTION	3
EVALUATION	7
DESIGN	7
RECRUITMENT AND RANDOMISATION	8
INSTRUMENTS	10
DATA COLLECTION AND TRAINING	12
BALANCE TESTS	12
RESULTS	14
CONTEXT	14
FIDELITY AND UPTAKE	17
NUMERACY AND LITERACY OUTCOMES	20
CONCLUSION	24
REFERENCES	26
APPENDIX	28

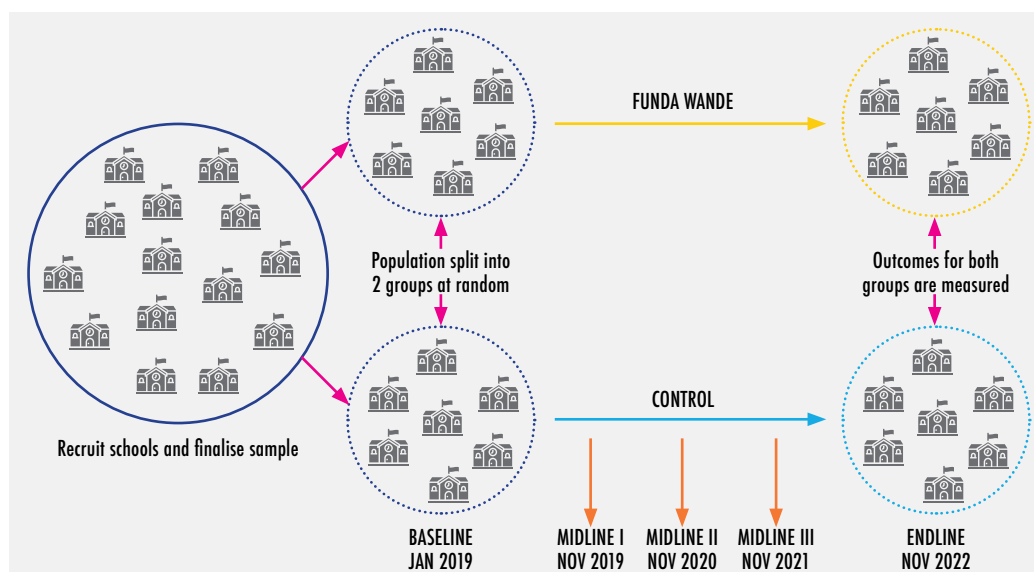
INTRODUCTION

Funda Wandé is a not-for-profit organization that aims to ensure that all learners in South Africa can read for meaning in their home language by the age of 10. They aim to achieve this through a combination of 1) the development of high quality open access materials and to equip teachers on how to teach literacy in African languages, 2) pre-service and in-service training, 3) research and advocacy and 4) iteratively experimenting with interventions to understand both ‘what works’ and how to scale. Partnering with the Eastern Cape Department of Education, Funda Wandé implemented a pilot intervention to support Foundation Phase literacy teachers in teaching reading for meaning over a four-year period (2019-2022). The essential components of the pilot intervention comprise the Funda Wandé learning and teaching support materials (LTSM) and the support of a team of expert coaches who are all experienced Foundation Phase educators. The coaches observe teachers in their classrooms, provide targeted advice on how to improve their practice, as well as providing model lessons with their students. Coaches visit each school an average of three times a month.

SALDRU at the University of Cape Town is conducting an external impact evaluation of the Funda Wandé Coaching Intervention. The primary aim of the evaluation is to assess the causal impact of Funda Wandé coaching on foundation phase learners’ ability to read with meaning. The impact evaluation uses a randomized control trial (RCT) with schools randomized into one of two arms – Funda Wandé and control – for a four-year period¹. Figure 1 provides a graphic representation of the RCT design.

¹ See Ardington and Meiring (2020) for full details on the evaluation design, selection and randomisation of schools.

Figure 1. Funda Wande Coaching Intervention Impact Evaluation design



The first year of the programme produced promising results with learners in Funda Wande schools outperforming their peers in control schools across all sub-domains of isiXhosa reading proficiency that could be reliably measured (Ardington and Meiring, 2020).

At the beginning of 2021, despite a year of COVID-19-related disruptions, learners in treatment schools had managed to maintain a statistically significant gap compared to learners in control schools across the range of EGRA sub-tasks (Ardington 2021). However, detailed analyses revealed that the point estimates on some of these gaps had diminished since the promising results of the first midline conducted at the end of 2019. For example, the gap between treatment and control learners in isolated word reading had decreased from about two words per minute at midline I to around one word per minute by midline II.

Unfortunately, school closures and rotational time-tabling did not end as hoped at the end of 2020, and 2021 saw serious ongoing disruptions to schooling. Reports from grade 1 teachers in the evaluation schools indicated that learners were attending school everyday in only 16 percent of schools in terms 1 and 2 (Ardington and Henry, 2021). This improved in the second half of the year but over one-third of schools had still not reverted to daily attendance by term 4.

This report focuses on the outcomes from the fourth and final wave of data collection and provides an opportunity to investigate whether the gaps between treatment and control learners widened or narrowed over the course of 2021.

2. COVID Impact on Intervention

COVID IMPACT ON INTERVENTION

The Funda Wande Coaching Intervention in the Eastern Cape relies heavily on face-to-face interaction between coaches and teachers, and was therefore heavily affected by the COVID-19 pandemic. Pertinent intervention components include teacher content training, in-class teacher coaching, and the completion of the 2-year Rhodes advanced certificate in teaching language in Foundation Phase by all Eastern Cape Department of Education (ECDoE) Department Heads, select ECDoE Subject Advisors, and Funda Wande coaches. In 2019, training was conducted in clusters per district and covered a variety of literacy topics from *Five big ideas in reading* to *Five CAPS reading activities*. Coaching included individual lesson observations, individual feedback sessions, as well as group sessions held after school.

Implementation in 2020

In 2020, the intervention planned to continue with the same design that had delivered promising early results in 2019. In term 1, all activities were executed as normal. The COVID-19 lockdown was announced in the last week of term 1, suspending all usual school and intervention activities. With the news of school closures, Funda Wande created a WhatsApp group for all parents for whom they were able to secure contact numbers. Although parental assistance with school work is historically very low in South Africa, the lockdown meant that many parents would be away from work which potentially would allow for increased school-work support in the home. Through the WhatsApp group, Funda Wande shared work activities with references to the Vula Bula anthologies. However, not all parents were contactable.

Once the schools began to re-open, the intervention activities had to change substantially. All cluster trainings were cancelled. The Rhodes course was amended from face-to-face into an online course which offered to teachers and to the broader ECDoE. Those in their second year of the Rhodes advanced certificate completed the qualification online. Individual lesson observations and feedback only recommenced in term 4 of 2022 and were limited in duration. No in-person group sessions were conducted in 2020. Online group sessions were held but data and connection problems were common. New homework books were created for learners in Term 4 in an attempt to counteract the negative effects of lost classroom time (see Figure 2).

Figure 2. Funda Wande COVID specific learning material



On-reopening, compliance with health, safety and social distancing requirements meant that the vast majority of schools needed to implement rotational timetabling. Over the course of 2020, between school closures and rotational timetabling, we estimate that learners lost around 60 percent of learning days in 2020 (Ardington et al., 2021). Learner exposure to the programme was effectively halved.

Implementation in 2021

The pandemic continued to affect in-person interactions and program exposure in 2021. There were no cluster trainings on material or content due to the COVID related restrictions on gatherings. In the first two months of the first term all coaching was conducted virtually, with no in-class support whatsoever. Limited-duration lesson observations and feedback resumed in March 2021. Group sessions were reinstated at the same time but were conducted in smaller groups than in 2019, diluting the exposure for individual teachers and resulted in content coverage taking a longer time.

In addition to limited classroom time for learners due to rotational timetabling, the alternative attendance system also meant that coaches only saw half the class on each visit. Many schools opted not to conduct Group Guided Reading at all due to social distancing guidelines.

TRAINING AND DATA COLLECTION

The endline fieldwork was conducted by iKapadata, the same company that conducted the first three waves of data collection. Fieldworker training for the fourth wave of data collection took place in Gqeberha from the 11th to 15th of October. Training was conducted by the Principal Investigator together with senior staff from iKapadata. Training consisted of five classroom days and one field day with trainees assessing learners in a pilot school. The classroom days included lectures, role play, simulated tests and inter-rater reliability (IRR) exercises. At the pilot school, all assessors completed at least four learner assessments. A total of 24 enumerator recruits were trained, of which four were hired as supervisors and 19 as enumerators.

All schools were contacted by telephone and emailed by iKapadata fieldwork staff members before the start of the actual data collection. The purpose of these calls was to reintroduce the study to the principal and relevant teachers and to gain permission to conduct the study at the school.

Data collection was conducted by four teams consisting of four fieldworkers and one supervisor between the 18th of October to the 30th of November.

Despite some challenges due to ongoing COVID-19 related rotational timetabling, the fieldwork was successfully completed. Table 1 presents the final sample of learners that were successfully assessed at each wave of fieldwork by cohort. At baseline, 10 learners from each school, of grade 1 (cohort A) and grade 2 (cohort B) were randomly selected for assessment. At midline II, 432 cohort A learners and 411 cohort B learners were assessed for the third time. At midline II, cohorts A and B should be in grades 3 and 4 respectively. However, 15 percent of cohort A are repeating grade 2 and 7 percent of cohort B are repeating grade 3.

Table 1. Learner cohorts by data collection wave

Fieldwork wave	Cohort A	Cohort B
Baseline (Jan/Feb 2019)	Grade 1 Term 1 n=590	Grade 2 Term 1 n=590
Midline I (Oct/Nov 2019)	Grade 1 Term 4 n=550	Grade 2 Term 4 n=559
Midline II (Feb/Mar 2021)	Grade 3 Term 1 n=432	Grade 4 Term 1 n=411
Endline (Oct/Nov 2021)	Grade 3 Term 4 n=461	Grade 4 Term 4 n=448

ATTRITION

The overall attrition rate between the baseline and midline I was six percent. COVID-19 had a considerable impact on attrition, with the overall attrition rate for midline II at 29 percent. However, by endline we managed to decrease the attrition rate to 22 percent. Table 2 shows the fieldwork outcomes

by cohort and wave of fieldwork. Higher attrition at midline II is driven by a dramatic increase in the proportion of learners who are no longer attending the school. Between midline I and midline II, 15 percent of learners left the school.

Table 2. Fieldwork outcomes by cohort and wave of fieldwork

	Cohort A			Cohort B		
	Midline I	Midline II	Endline	Midline I	Midline II	Endline
Assessed	93%	73%	79%	95%	70%	76%
School refused	2%	3%	3%	2%	5%	3%
Learner no longer at school	2%	18%	16%	<1%	15%	18%
Learner absent	2%	6%	2%	2%	10%	3%
Other	1%			<1%		

Table 3 presents the attrition rate by treatment status, wave of fieldwork and cohort. The absolute difference in the attrition rate between treatment and control groups is shown in the final row of the

table. Attrition was slightly higher for the treatment group at midline I and for the control group at midline II. Endline attrition rates were very similar for treatment and control.

Table 3. Rate of attrition by wave and cohort

	Cohort A			Cohort B		
	Midline I	Midline II	Endline	Midline I	Midline II	Endline
Treatment	8%	25%	21%	6%	29%	24%
Control	5%	29%	23%	4%	32%	24%
Absolute difference	3%	4%	2%	2%	3%	0%

BALANCE BETWEEN TREATMENT AND CONTROL

Baseline equivalence of the full sample was guaranteed through the random assignment of schools to treatment and control using statistical software. Checks to assess attrition suggest that any differential attrition between learners in treatment and control schools does not lead to an imbalanced sample at endline. Firstly, the differences in attrition between treatment and control shown in Table 3 are small and not statistically significant. Second, even with the substantial increase in attrition compared to midline I data collection, the overall attrition (23 percent) and minimal differential attrition (0.4 percent) fall within the What Works Clearinghouse (WWC) (2020) conservative limits for low expected bias, and therefore meets the highest possible WWC standards. Furthermore, appendix Table A1 shows that non-attriting learners are still balanced in terms of a range of baseline assessment scores, observable learner characteristics and household assets. The differences between treatment and control on all variables are within the limits to satisfy baseline equivalence. The low differential attrition and the baseline equivalence of the two groups of non-attriting learners mean that programme impacts can still be reliably estimated.

EARLY GRADE READING ASSESSMENTS (EGRA)

During each wave fieldworkers administered a one-on-one Early Grade Reading Assessment (EGRA) with each learner. The assessment components differed slightly by grade to take into account increased reading proficiency over time. Table 4 presents the EGRA sub-tasks that we include in the analysis and indicates for each cohort, the waves in which the specific sub-task was administered.

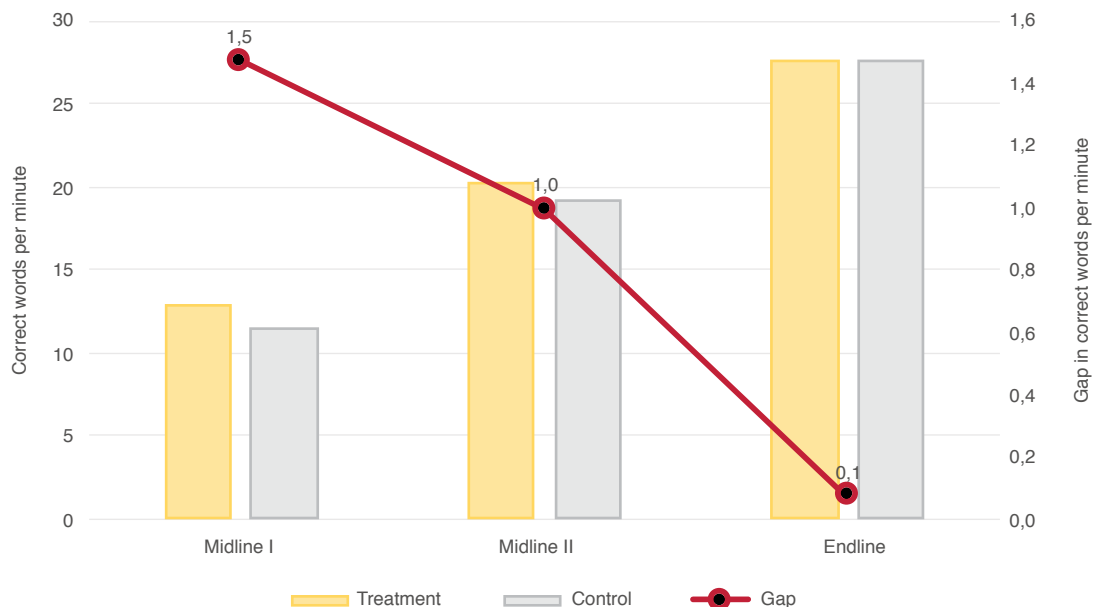
Table 4. EGRA sub-tasks by cohort and fieldwork wave

	Cohort A			Cohort B		
	Midline I	Midline II	Endline	Midline I	Midline II	Endline
Listening comprehension						
Phonemics awareness						
Letter sounds						
Digraphs Trigraphs						
Isolated words						
Oral reading fluency (ORF)						
ORF comprehension						
ORF Passage 2						
ORF Comprehension 2						

RESULTS

We use the endline data collected in term 4 of 2021 to investigate whether the gaps between treatment and control learners widened or narrowed over the course of 2021. We begin with a visual inspection of the gap between treatment and control learners on the first oral reading fluency task. In order to hold the composition of the sample constant across time, we restrict the analysis to learners who were assessed in all four waves. The bars in Figure 3 present mean oral reading fluency at midline I, midline II and endline separately for treatment (yellow) and control (grey). The figure also includes a red line showing the evolution of the gap between treatment and control learners over time. For oral reading fluency, we can clearly see that the positive gains initially made pre-COVID were eroded over the course of 2020 and 2021. By endline the treatment and control groups had very similar oral reading fluency.

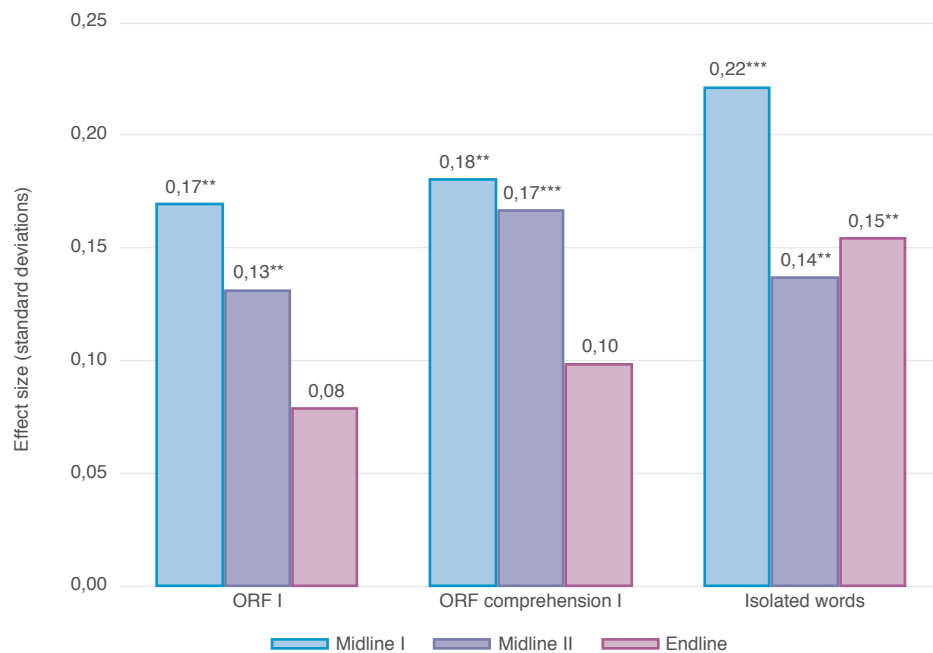
Figure 3. Oral reading fluency by wave and treatment status



We investigate whether these gaps are statistically significant using regression analysis that allows us to control for baseline measures of reading proficiency, learner characteristics (gender, age, height-for-age z-score) and household assets. Details of the regression models used are provided in the appendix and a full set of results are shown in appendix Table A1. We examine all four waves of data to analyse whether there were differential impacts between baseline and midline I, between midline I and midline II, and finally between midline II and endline. Figure 4 summarises the gaps, or treatment effects, at each wave for EGRA tasks that were completed by all learners at midline I, midline II and endline. For comparability purposes, all treatment effects are shown in standard deviations. The asterisks indicate that the treatment effect is significant at the 1% (***) and 5% (**) levels. Starting with the first oral reading

fluency task, we see that treatment learners had average scores at midline I that were 0.17 standard deviations higher than their peers in control schools. Between midline I and midline II, there were no further improvements for treatment over control learners. In fact, the gap between treatment and control learners narrowed over 2020 to 0.13 standard deviations. Nevertheless, treatment learners still significantly outperformed control learners. Pre-COVID gains were further eroded in 2021 with the gap diminishing to 0.08 standard deviations with the result that the treatment group now look statistically similar to the control group. We see a similar erosion of early programme gains for reading comprehension and isolated word reading although the difference between treatment and control learners remains significant at endline for isolated word reading.

Figure 4. Standardized treatment effects at midline I, midline II and endline



We move from an analysis of whether treatment impacts differ over time to a simple focus on outcomes in the fourth wave of data collection at the end of 2021. This widens our scope to all EGRA tasks administered at endline and increases our analytical sample as there was negative attrition between the third and fourth waves of fieldwork. Table 5 provides a summary of the estimates of the endline treatment effects (equation 2 in Appendix A2). Looking at the first column, we see that the gap between treatment

and control learners in overall reading proficiency² was small (0.1 standard deviations) and no longer statistically significant by the end of 2021. Across the sub-tasks, all point estimates are positive. However, they are mostly small and not statistically significant at the conventional 5 percent level. The exception is listening comprehension where grade 3 learners in treatment schools score four percentage points higher on average than their peers in control schools.

Table 5. Endline treatment impacts

	Cohorts A and B						Cohort A			
	Composite score (SD)	ORF	ORF comprehension	ORF Passage 2	ORF Comprehension 2	Isolated words	Listening comprehension	Phonemic awareness	Letter sounds	Digraphs Trigraphs
Treatment	0.101	0.939	0.0332	1.659	0.0408*	1.798*	0.0389**	0.00613	5.286*	3.574
	(0.0645)	(1.252)	(0.0228)	(1.030)	(0.0233)	(0.926)	(0.0182)	(0.0184)	(2.815)	(2.329)
Observations	881	890	884	889	888	895	457	457	457	457
R-squared	0.473	0.463	0.384	0.477	0.381	0.471	0.264	0.284	0.266	0.302

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

² We used Principal Components Analysis (PCA) to reduce the data on the six sub-tasks that were administered to both cohorts A and B into a single variation that captures 80 percent of the common variation among them. The PCA was restricted to learners in control schools so that the weightings of each sub-task reflect those in 'business as usual' environments. These weightings were then applied to all learners. The composite score is measured in standard deviations of control learner performance.

CONCLUSION

COVID-related schooling disruptions have had a devastating impact on learning (Ardington et al., 2021) and the more targeted individualised learning promoted and supported by Funda Wande could prove crucial in helping teachers remediate these gaps. However, the extent of the disruptions in 2020 and 2021 appear to have seriously hampered the ability of Funda Wande coaches to support learning and remediation efforts. COVID health and safety regulations substantially curtailed face-to-face interactions between teachers and coaches, both on and off site. Ongoing rotational timetabling throughout 2020 and 2021 limited learner exposure to the Funda Wande intervention. In this context, there appears to have been no additional advantage to being in a Funda Wande school during 2020 and 2021. It is disheartening, but perhaps to be expected, that learners in treatment schools were not able to maintain their significant advantage over learners in control schools that we observed after the first year of the programme.

REFERENCES

- Ardington, C. and Meiring, T. (2020) Impact Evaluation of Funda Wande Coaching Intervention Midline Findings: Midline I. SALDRU report. University of Cape Town
- Ardington, C. and Henry, J. (2021) Funda Wande Eastern Cape Evaluation: Midline Findings. SALDRU report. University of Cape Town
- Ardington, C. (2021) Impact Evaluation of Funda Wande Coaching Intervention Midline Findings: Midline II. SALDRU report. University of Cape Town
- Ardington, C., Wills, G. and Kotze, J. (2021) COVID learning losses: early grade reading in South Africa. International Journal of Educational Development
- What Works Clearinghouse. (2020) What Works Clearinghouse Standards Handbook, Version 4.1. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. This report is available on the What Works Clearinghouse website at <https://ies.ed.gov/ncee/wwc/handbooks>

APPENDIX

APPENDIX A1. ENDLINE BALANCE

Table A1. Endline balance tests and effect sizes

	Treatment		Control		p-value	Pooled s.d.	Effect size
	Mean	s.e.	Mean	s.e.			
Common tasks							
Letter Sounds per minute	18.1	20.3	17.8	19.0	0.880	19.7	0.01
Digraphs and Trigraphs per minute	5.9	12.0	4.9	10.9	0.378	11.5	0.09
Productive Listening Comprehension	3.1	1.5	3.1	1.5	0.975	1.5	0.00
Receptive Listening Comprehension	9.6	0.7	9.5	1.0	0.111	0.9	0.16
Phonemic Awareness	2.9	2.6	3.0	2.5	0.798	2.5	0.03
Expressive Vocabulary	11.3	3.7	11.7	4.0	0.392	3.9	0.09
Write letters	2.7	1.9	2.9	1.8	0.186	1.9	0.12
Grade 1 only tasks							
Word Choice	1.8	1.8	2.0	1.6	0.184	1.7	0.15
Rapid Automatized Naming	37.0	11.8	38.4	11.7	0.231	11.7	0.12
Write your name	4.6	1.0	4.6	0.8	0.814	0.9	0.03
Copy a word	4.4	1.3	4.3	1.3	0.577	1.3	0.07
Grade 2 only tasks							
CVCV Words per minute	11.4	12.9	10.2	12.2	0.537	12.6	0.09
Familiar Words per minute	7.9	8.9	7.1	8.8	0.564	8.8	0.09
Oral Reading Fluency	8.4	10.0	8.0	9.4	0.757	9.7	0.05
Reading Comprehension	3.5	3.8	3.4	3.6	0.860	3.7	0.03
Vocabulary	3.3	2.4	3.4	2.3	0.940	2.3	0.01
Sentence Comprehension	5.0	4.5	4.8	4.3	0.785	4.4	0.04
Write words	14.5	6.8	15.1	6.2	0.542	6.5	0.10
Learner characteristics							
Grade 1	0.5	0.5	0.5	0.5	0.585	0.5	0.02
Grade 2	0.5	0.5	0.5	0.5	0.585	0.5	0.02
Female	0.5	0.5	0.5	0.5	0.144	0.5	0.07
Age in months	6.5	0.9	6.4	0.9	0.224	0.9	0.10
Height for age z-score	-0.3	1.0	-0.3	1.0	0.790	1.0	0.02
Household assets							
Books other than schoolbooks to read at home	0.3	0.5	0.4	0.5	0.023	0.5	0.18
Radio	0.7	0.5	0.7	0.5	0.919	0.5	0.01
Television	1.0	0.2	0.9	0.2	0.430	0.2	0.05
Computer	0.3	0.5	0.3	0.5	0.608	0.5	0.04
Toilet	0.6	0.5	0.6	0.5	0.508	0.5	0.06
Vehicle	0.5	0.5	0.5	0.5	0.764	0.5	0.03

Notes: Effect sizes are calculated as the difference in means between the treatment and control groups, divided by the pooled standard deviation for the variable. If the effect sizes are 0.05 or less in absolute value, the two groups are considered equivalent on that dimension. When effect sizes are in the range between 0.05 and 0.25, the baseline measures should be included as controls in the model estimating programme effects to satisfy equivalence. Variables for which such adjustments are required include 10 of the 18 sub-tasks, learner age, learner gender, whether they have books to read at home, and whether their household owns a television or has an inside toilet. These variables are therefore added as controls to satisfy equivalence between the two groups, and not only to improve the precision of the estimates of programme impact. No effect size is greater than 0.25 in absolute value - the level at which the samples are considered to be not equivalent.

APPENDIX A2: REGRESSION EQUATIONS AND DETAILED RESULTS

We begin by examining all four waves of data to analyse whether there were differential impacts between baseline and midline I, between midline I and midline II, and finally between midline II and endline. Specifically, we estimate equations of the following form:

$$Y_{igsat} = \beta_0 + \beta_1 (Treatment)_s + \beta_2 (Midline II) + \beta_3 (Treatment \times Midline II) + \beta_4 (Endline) + \beta_5 (Treatment \times Endline) + X'_{igsd0}T + \mu_d + \varepsilon_{igsat} \text{ for } t \in (1,2,3) \quad (1)$$

Where the outcome measure, Y_{igsat} is a measure of reading proficiency for learner i in grade g of school s in district d at time t (where $t=0$ is baseline, $t=1$ is midline I, $t=2$ is midline II and $t=3$ is endline); $(Treatment)$ is the dummy variable indicating the treatment status for school s ; X'_{igsd0} is a vector of baseline controls, μ_d are the specific school district (or strata) fixed effects; and ε_{igsat} is the error term clustered at the school level. The parameters of interest are β_1 , the average treatment effect at midline I; β_3 , the additional impact of treatment between midline I and midline II and; β_5 , the additional impact of treatment between midline II and endline. Negative values of β_3 and β_5 will indicate a narrowing of the previous gap between treatment and control learners while a positive value would indicate that there were additional improvements after the first midline.

Baseline controls in all regressions include each of the baseline measures of reading proficiency, learner characteristics (gender, age, height-for-age z-score) and household assets. In the cases where learners had missing data on a certain dimension of the control variables, a missing value was imputed and a separate dummy variable was included to indicate missingness as a control. Standard errors are clustered at the level of randomisation (i.e. schools).

The sample for all regressions was restricted to the balanced panel, i.e. learners who were assessed in all four waves.

We also provide simple estimates of the treatment effect at endline from regressions of the following form:

$$Y_{igsat} = \beta_0 + \beta_1 (Treatment)_s + X'_{igsd0}T + \mu_d + \varepsilon_{igsat} \quad \text{for } t \in (3) \quad (2)$$

These regressions use the full wave 4 sample.

Table A2. Treatment impacts at midline I, midline II and endline

	Cohorts A and B			Cohort A				Cohort B	
	ORF I	ORF comprehension I	Isolated words	Listening comprehension	Phonemic awareness	Letter sounds	Digraphs Trigraphs	ORF II	ORF comprehension II
Treatment	2.356***	0.0575**	2.571***	0.0370	0.0627**	4.111*	3.797**	1.616	0.0621**
	(0.879)	(0.0222)	(0.697)	(0.0249)	(0.0236)	(2.333)	(1.748)	(1.032)	(0.0289)
Treatment x Midline II	-0.194	-0.00272	-0.719	-0.00271	-0.0402	0.0837	1.427	0.0599	-0.0150
	(0.710)	(0.0185)	(0.498)	(0.0250)	(0.0243)	(1.714)	(1.550)	(1.000)	(0.0263)
Treatment x Endline	-0.943	-0.0252	-0.349	-0.0130	-0.0606**	0.659	0.294	0.207	-0.0180
	(0.965)	(0.0209)	(0.688)	(0.0221)	(0.0256)	(2.359)	(2.144)	(1.117)	(0.0307)
Midline II	7.636***	0.168***	6.646***	0.0891***	0.101***	9.012***	10.15***	6.331***	0.141***
	(0.445)	(0.0108)	(0.337)	(0.0189)	(0.0168)	(1.152)	(0.974)	(0.743)	(0.0198)
Endline	15.86***	0.243***	11.79***	0.148***	0.173***	20.23***	19.66***	13.18***	0.222***
	(0.668)	(0.0125)	(0.440)	(0.0170)	(0.0180)	(1.660)	(1.364)	(0.779)	(0.0259)
Observations	2,190	2,190	2,190	1,128	1,128	1,128	1,128	1,062	1,062
R-squared	0.577	0.502	0.593	0.329	0.350	0.388	0.429	0.599	0.487

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

Level 3, School of Economic Building, Middle Campus, University of Cape Town, Rondebosch, 7700.