AFEP | Africa Fellows in Education Program

GLOBAL EDUCATION ANALYTICS INSTITUTE
Putting evidence at the heart of education policy

AFEP WORKING PAPER

The Long-Term Impact of an Early Grade Reading Intervention on Learning: Evidence from South Africa's Early Grade Reading Study

Zamangwe Zwane

Department of Basic Education, South Africa

April 2026

geai.io

A GEAI program, implemented in partnership with pep partnership for economic policy

Acknowledgements

This work was carried out with financial and scientific support from the Partnership for Economic Policy (PEP) and the Global Education Analytics Institute (GEAI) through funding provided by the Yidan Foundation. The author is grateful to Sahawal Alidou, who served as a mentor through PEP, for his guidance and insightful feedback throughout the development of this work. Sincere thanks are also extended to Stephen Taylor and the Department of Basic Education for providing access to the administrative and assessment data used in this study, for assisting with the matching of intervention learners to administrative records, and for valuable comments on earlier drafts. The views expressed do not necessarily reflect those of the Department of Basic Education.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Partnership for Economic Policy (PEP) or the Global Education Analytics Institute (GEAI).

The Long-Term Impact of an Early Grade Reading Intervention on Learning: Evidence from South Africa's Early Grade Reading Study

Zamangwe Zwane*

Department of Basic Education, South Africa

April 2026

Abstract

Programme models intended to improve learning at scale have increased in low-and middle-income countries. Although some of these programmes have led to gains in learning outcomes, very little research has been conducted on the extent to which the effects of these scalable models and programmes persist, especially beyond initial improvements in early grades. Using evidence from South Africa's Early Grade Reading Study, a randomised control trial with three treatment arms implemented from 2015-2017, we assess whether gains in foundational reading skills achieved in the second or third year of schooling translate into better learning outcomes in secondary/high school education. By means of an estimation equation employed in a previous evaluation of the study, we evaluate administrative data to show a sustained impact of the intervention, with treated learners outperforming their control group peers by 0.12 standard deviation in overall high-school subject performance in Grades 8 and 9. The results from this study support the argument that investment in foundational skills during early grades can lead to long-term improvements in learner outcomes, particularly through the coaching intervention approach implemented in the Early Grade Reading Study.

Keywords: Long-term outcomes, impact evaluation, early grade reading, high-school outcomes

* Department of Basic Education, Republic of South Africa. Email: zamiizwane@gmail.com

1 Introduction

Increasingly, developing countries are prioritizing foundational reading skills with the expectation that improvements in that area will stimulate the development of new cognitive abilities and result in better life outcomes for learners (Kim, Lee & Zuilkowski, 2020; Rodriguez, 2022; Stern et al., 2024). Some of the programmes that have been developed to improve learning at scale in low-and middle-income countries include:

- The Gauteng Primary Language and Mathematics Strategy was implemented in South Africa in 2011 as a government-funded programme that aimed to improve learner outcomes in Home Language (mostly African languages), English as First Additional Language, and Mathematics in 1,000 Quintile 1, 2, and 3 schools in which Grade 1-3 learners were enrolled.[1] The design of the intervention made use of a "best practice" model for systemic change, including a combination of lesson plans, quality learning materials, as well as both just-in-time training and one-on-one instructional coaching for more than 8,000 teachers (Fleisch & Alsofrom, 2022). The effectiveness of the Gauteng Primary Language and Mathematics Strategy was measured using a regression discontinuity design which compared changes in school average scores between comparable groups of schools (Fleisch et al., 2016; University of the Witwatersrand & Fleisch, 2018). The evaluation results revealed that the instructional-coaching model was very effective in improving learning outcomes and that a model of system-wide coaching as part of a structured pedagogic programme in primary schools was financially viable. Although the strategy led to large gains in both reading and mathematics, it did not survive the political transition from one provincial minister of education to another and ended in 2014.

- The Reading Catch-Up Programme was implemented in South Africa in 2012 and used the embedded instructional-coaching model with lesson plans and quality materials. The Catch-Up Programme was aimed at improving learning outcomes in English as a first additional language among Grade 4 learners and made use of the same structed lesson plans, storybooks, and coaching as the Gauteng Primary Language and Mathematics Strategy. An impact evaluation of the programme was conducted in 2014, and the results showed that, although performance improved in both control and intervention schools, intervention schools performed slightly better (Fleisch et al., 2017). Slight gains were noted in spelling and grammar performance, but learners did not improve enough in English literacy to allow them to deal with the demands of the Grade 4 curriculum.

[1] The 1996 South African Schools Act required the government to fund public schools, and the National Norms and Standards for School Funding policy introduced a quintile-ranking mechanism to ensure equity. Quintile 1 schools cater to the poorest 20% of learners, Quintile 2 schools to the next poorest 20% of learners, and so on. Quintile 1, 2, and 3 schools are classified as "no-fee" and receive larger state subsidies (Department of Basic Education, 2004).

- Between 2016 and 2017, Zambia's Ministry of Education, with UNICEF support, piloted a remedial Catch-Up Programme based on Teaching at the Right Level methodology (UNICEF, 2022). The programme was targeted at learners in Grades 3, 4, and 5 in eighty schools across four districts in the Eastern and Southern Provinces of Zambia. Learners were grouped according to literacy and numeracy levels rather than by age or grade and received one hour per day of interactive lessons. The programme was later upscaled to all 1,877 schools in Zambia's Eastern and Southern provinces. In 2019, results from the scaled-up programme showed an increase in the number of Grade 3-5 learners who could read a simple paragraph, from 68,892 learners at the beginning of the year to about 110,228 learners at the end of the year; in 2020, the programme increased literacy and numeracy skills among 239,663 students in Grades 3-5.

Since 2008, the ministries of Education in Liberia and Kenya have worked to improve reading outcomes across primary school learners, transitioning from small-scale pilot programmes to the large-scale implementation of early-grade reading interventions {Citation}. Liberia developed the Early Grade Reading Assessment Plus programme while Kenya inaugurated the Primary Math and Reading initiative with funding from the United States Agency for International Development. Liberia's Early Grade Reading Assessment Plus programme (2008-2011) had two treatment arms and a control group; the control group received the intervention after endline assessments. Each of the treatment groups and the control group included sixty schools in which Grade 2 and 3 learners were enrolled.

The "light" treatment consisted of support for classroom-based assessment as well as the dissemination of students' results to parents and community members (Gove, Poole and Benjamin, 2017). In the full treatment, teachers were trained by instructional coaches for ten days over two semesters and received instructional materials and a manual with scripted lesson plans. A difference-in-difference strategy was employed to evaluate the effect of Early Grade Reading Assessment Plus. This evaluation showed that the full treatment increased student achievement for every Early Grade Reading Assessment subtask with an overall effect size of 0.79 standard deviation, ranging from 0.39 standard deviation to 1.23 standard deviations. In 2011, based on these results, Liberia implemented a similar large-scale programme across Grades 1-3 in 1,200 schools (Gove, Poole and Benjamin, 2017). Letter-naming and phonemic awareness were the only two skills that showed significant improvements relative to the control group.

Kenya's Primary Math and Reading initiative was designed in response to a study that revealed very low reading performance in several parts of Kenya. The focus of the programme was professional development for teachers to support large-scale instructional change (Gove, Poole and Benjamin, 2017). The Primary Math and Reading pilot was implemented by Kenya's Ministry of Education between 2011 and 2014 across 1,384 schools. Under the Primary Math and Reading initiative, students received English and Kiswahili textbooks, and teachers were

given a corresponding teaching guide. As a third component to the intervention, teachers were trained for seven days a year, during which they observed and practiced a model mini lesson. The final component of the intervention was instructional support through Teachers Advisory Centre tutors who focused on improving teaching and providing in-classroom instructional support (Gove, Poole and Benjamin, 2017). The Primary Math and Reading programme produced effect sizes between 0.3 and 0.8 standard deviation in reading. The intervention resulted in moderate-to-large effects for all subjects, including English, Kiswahili, Mathematics, and mother tongue. The Tusome programme (2014-2019) was a large-scale primary math and reading programme implemented in over 23,500 schools (Gove, Poole and Benjamin, 2017) and has been evaluated several times, though not always through randomized control trials. The results of an evaluation of the Tusome programme by (Piper et al., 2018) gave encouraging evidence that countries facing education quality challenges can successfully implement large-scale, coherent instructional reforms through existing government systems.

Although gains in learning outcomes in the early grades have been demonstrated from some of the interventions mentioned above, little research has focused on the extent to which these scalable models and programmes have had lasting effects on learning outcomes, especially beyond initial improvements in early grades (Stern et al., 2023). Research has shown that gains in learners' reading fluency through exposure to print improves cognitive skills, including oral language, basic reading skills, spelling, content/declarative knowledge, and vocabulary. Learners who spend more time engaging in reading develop more advanced vocabularies and stronger language comprehension skills (Sparks, Patton and Murdoch, 2013). The exposure of these learners to print through frequent reading increasingly widens an academic-achievement gap known as the Matthew Effect.[2]

In 2008, the Department of Basic Education in South Africa developed a National Reading Strategy, using resources and techniques to improve learners' reading competencies through increased access to books and additional teacher support (Mawoyo & Vally, 2020). One of the interventions implemented to support this strategy was the Early Grade Reading Study. The approach employed in the Gauteng Primary Language and Mathematics Strategy study, which included lesson plans, materials and instructional coaching, was reintroduced in the form of a collaboration between researchers from the South African Department of Basic Education and the University of Witwatersrand.

The Early Grade Reading Study was designed to test a basic model of instructional coaching combined with daily lesson plans and quality educational materials and allowed a comparative

[2] The Matthew Effect is a Biblical reference in which the rich get richer, and the poor get poorer; the very children who can read and have good vocabularies will read more, learn more word meanings, and, as a result, read better (Stanovich, 1986). Children with weaker vocabularies who read slowly will read less causing them to have slower development of vocabulary, which as a result prevents further growth in reading ability (Zenex Foundation, 2022).

analysis of three distinct intervention models: lesson plans, educational materials, and training that included embedded coaching (Taylor et al., 2025). After two years of implementation, the basic model of instructional coaching embedded in a structured pedagogic programme showed the most consistent gains. The evaluation found that it had statistically significant positive effects on all Home Language literacy measures and on English sub-tests of 0.25 standard deviation (Department of Basic Education et al., 2018). A further impact evaluation, conducted a year after the completion of the program in 2018, revealed that only teachers who received coaching maintained improved teaching techniques (Cilliers et al., 2022; Department of Basic Education et al., 2016). The impact evaluation also indicated the magnitude of effects on learning for students taught by teachers one year after program participation with 0.11 standard deviation for the training arm and 0.19 standard deviation for the coaching arm (Cilliers et al., 2022; Department of Basic Education et al., 2016). While sustained gains for learners were observed, the impact on the second cohort had halved. A further wave of evaluation data was collected in 2023 following the original cohort of learners then in Grade 7, the final year of primary school, four years after the students participated in the programme. Results showed a sustained impact of the coaching intervention on Setswana Home Language literacy of 0.17 standard deviation as well as an improved grade attainment: almost 10% more students reached Grade 7 by the seventh year of primary school (Department of Basic Education et al., 2024; Stern et al., 2024). The findings also showed an impact on the reduction of non-readers and an expansion of higher-order skills such as English written comprehension (Reading a comprehension for themselves and answering questions based on text that was read).

Previous evaluations of Early Grade Reading Study I have shown that the coaching-based, structured-pedagogy intervention produced sizeable gains in Setswana Home Language literacy in Grades 1-3. Some of these gains persisted through Grade 7, including improved grade attainment and reduced numbers of non-readers. What is still unknown is whether the early improvements in foundational reading skills led to better learning outcomes in secondary school or supported timely progression through the initial years of high school. We have linked the original Early Grade Reading Study learner cohort to School Based Assessment data from the Data Driven Districts system for Grades 8 and 9 in 2023, thereby providing the first evidence on the long-term impact of Early Grade Reading Study I on high school performance in South Africa.

This paper contributes to the literature substantively by providing new evidence on the long-term effects of an early-grade reading intervention in a low-and middle-income country context. Methodologically, it shows how administrative education data, specifically School Based Assessment data, can be used to track learners and conduct long term follow-up analyses of randomized controlled trials.

More specifically, the research objectives were:

1) To assess the persistence of learning gains achieved through the Early Grade Reading Study in later years of schooling.

2) To explore the impact of Early Grade Reading Study on educational outcomes such as grade repetition and the reduction of non-readers.

These objectives were measured in the first two years of high school,[3] Grades 8 and 9.

1.1 Research Context

South Africa has an unequal education system in which 93% of schools are public, and its overall quality is poor. Approximately 83% of learners attend no-fee schools in the poorer parts of the system and the remaining minority attend fee-paying public schools and private schools. (Department of Basic Education, 2020a). South Africa has a legacy of colonialism and apartheid, both resulting in deeply entrenched race-based systemic inequalities. Understanding and addressing the reasons for these continued inequalities and poor learning outcomes have been at the forefront for government, private sector actors, educators, and academics in the country.

Learning outcomes in the South African education system are very poor, with most learners performing two or three grades below required grade level (Department of Basic Education, 2020b). A 2007 meta-evaluation by the Zenex Foundation of thirty-four projects conducted by revealed that learners were being promoted from grade to grade without having mastered the content and foundational competencies of the previous grades (Zenex Foundation, 2022).This resulted in cognitive learning backlogs that prevented the development of more complex competencies. These were exacerbated by a change from poor learning outcomes in Home Language to English as a medium of instruction, which is used by 90% of schools from Grade 4 onwards. Learners in South Africa score poorly on both national and international assessments: results from local Grade 3 systemic evaluations showed that a national average of 20% of learners were at the required reading literacy level, and, in the Progress in International Reading Literacy Study (PIRLS) 2021, pandemic notwithstanding, only 19% of learners reached the low international benchmark (understanding literal information in texts) (Department of Basic Education,, 2022; Mullis et al., 2023). These results suggests that large gaps remain in early literacy development in the Foundation Phase (through Grade 3). This learning crisis is common to education systems in low-and middle-income countries in which 53% of Grade 4 learners cannot read for meaning (Mawoyo and Vally, 2020). In Sub-Saharan Africa, an average of 60% of learners are unable to correctly read a single word of a prose passage (USAID, 2015).

Teachers' pedagogical-content knowledge in language should be strengthened because gaps exist between the language skills of students entering initial teacher education and the skill teach-

[3] In South African, primary School is from Grades 1-7 and High School is from Grades 8-12.

ers need to be effective in schools (Roberts et al., 2024).The 2013 study by the Southern Africa Consortium for Monitoring Educational Quality showed that 79% of Grade 6 mathematics teachers could not score 60% on a primary school mathematics test (Department of Basic Education, 2017). A study by Bowie, Venkat, and Askew (2019) tested 488 first-year B.Ed. students and 282 final-year B.Ed. students from three South African Universities on primary-school-level outcomes; the average score for the first year B.Ed. students was 52% and 54% for the final year B.Ed. students. These results show that at least half of primary school teachers lack the primary-school-level content knowledge needed to teach their subjects (Spaull & Taylor, 2022). Given the gaps in their knowledge and skills, teachers face challenges in covering grade-level curriculum and dealing with learning backlogs (Zenex Foundation, 2022).

Learners with low attainment in the early grades tend to accumulate larger learning backlogs over time, particularly when teachers are not adequately equipped. All these challenges contribute to learner dropout, particularly in the first two years of high school when compulsory schooling ends at the age of 15 (Grades 8 and 9). Table 1 gives a brief picture of the dropout rates of high school learners in South Africa and shows the dropout and survival rates for three different age cohorts that were part of a survey called the General Household Survey: those born during 1979-1981 (and surveyed between 2003-2005), those born during 1987-1989 (and surveyed between 2011-2013), and those born during 1996-1998 (surveyed between 2020-2022).[4] The survival rate in the table refers to the percentage of a cohort of learners enrolled in the first grade in a specific school year who were expected to reach a specific grade (Department of Basic Education, 2024). Across all three age cohorts, the results showed that dropout rates were low in the primary school Grades and increased with each Grade, peaking in the high school grades. In the most recent survey of the cohort between 2020 and 2022, for example, 4.5% of those who reported having completed Grade 8 completed no further education and were therefore assumed to have dropped out of school. This dropout rate was almost double that of those who dropped out of school at the end of primary school (Grade 7). A gradual increase in high school dropout rate can be observed from Grade 8 through to Grade 12 with a peak after Grade 11.

Figure 1 shows grade-repetition statistics of learners from Grade R (pre-primary) to Grade 12 surveyed between 2009 and 2022. The Figure shows that grade-repetition rates have typically been higher in secondary school than in primary school, peaking in grades 10 and 11 at almost 20% (Department of Basic Education, 2024). The lower grade-repetition rates observed in 2021 compared to the other years resulted from relaxed promotion requirements in response to pandemic-related disruptions in the 2020 school year.

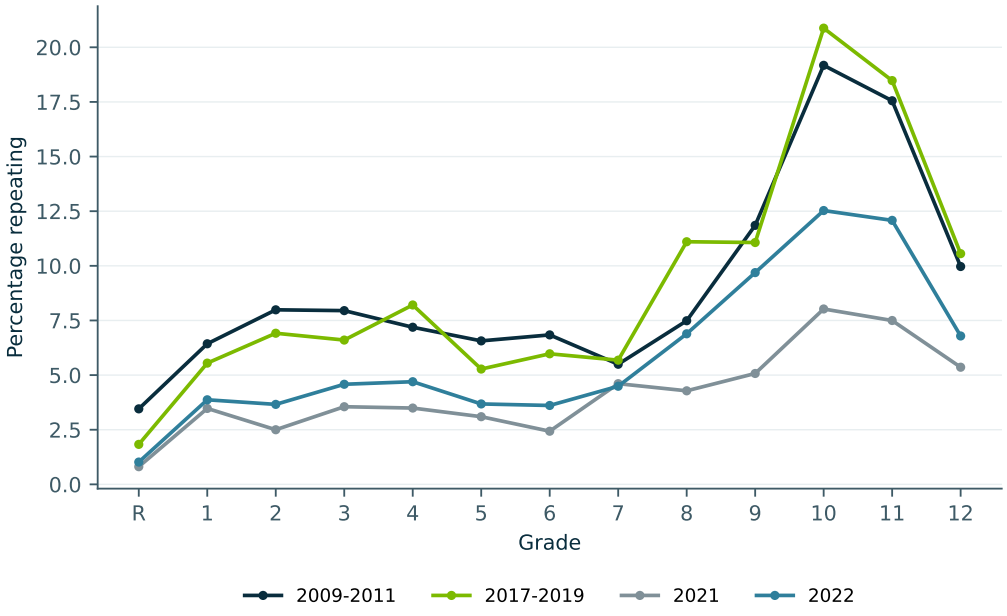
[4] The GHS is an annual survey that measures aspects of South African households' living circumstances and is conducted by statistics South Africa in around 22,000 households (Department of Basic Education, 2024). The survey is one of the key sources of information on school nutrition programmes, grade repetition, age-grade enrolment rates, and a number of other indicators of interest in the education sector. The survey was first compiled in 2002 and has been continued annually since then.

Table 1: Survival and Dropout Rates by Grade

2003-2005 (Born 1979-1981)			2011-2013 (Born 1987-1989)		2020-2022 (Born 1996-1998)	
	Survival Rate	Percentage dropping out with this grade attained	Survival Rate	Percentage dropping out with this grade attained	Survival Rate	Percentage dropping out with this grade attained
Total cohort	**100%**		**100%**		**100%**	
No schooling		1.8%		0.7%		0.6%
Grade 1	98%	0.3%	99.3%	0.4%	99.4%	0.2%
Grade 2	98%	0.5%	98.9%	0.5%	99.2%	0.1%
Grade 3	97%	0.9%	98.4%	0.4%	99.1%	0.3%
Grade 4	97%	1.5%	98.0%	0.7%	98.9%	0.4%
Grade 5	95%	1.9%	97.3%	1.0%	98.5%	1.1%
Grade 6	93%	3.2%	96.3%	1.5%	97.4%	1.6%
Grade 7	90%	5.5%	94.9%	3.0%	95.9%	2.8%
Grade 8	85%	7.9%	92.1%	5.1%	93.2%	4.5%
Grade 9	79%	11.2%	87.3%	8.9%	89.0%	9.4%
Grade 10	70%	18.2%	79.6%	17.3%	80.6%	10.4%
Grade 11	57%	28.1%	65.8%	26.7%	72.2%	21.3%
Grade 12	41%		48.2%		56.8%	

Notes: Source: Department of Basic Education, Republic of South Africa, General Household Survey (GHS): Focus on Schooling 2022 Report (Department of Basic Education, 2024).

Figure 1: Percentage of Learners Repeating Current Grade



Notes: ***Source:*** *Source: Department of Basic Education, Republic of South Africa, General Household Survey (GHS): Focus on Schooling 2022 Report* (Department of Basic Education, 2024).

A review of reading from Grade R to university learners found a concerning gap between learners' actual reading levels and grade-appropriate reading levels by the time learners reached the Senior Phase (Grades 7-9) (Zenex Foundation, 2022). The average age of Grade 8 learners was 14.4 years; however, in the READ annual report of 1999[5] it was found that the English (as a second language) reading level of Grade 8 learners in rural areas was equivalent to those of children aged 7.6 years who were Grade 2 learners. This explains the increasing levels of learner dropout. To understand which early-learning skills and intervention approaches have the greatest impact on later educational outcomes, including in reducing learner dropout and grade repetition, it is important to measure the long-term effects of effective early grade reading interventions to inform the planning and design of subsequent interventions or scale ups(Spaull & Taylor, 2022).

1.2 Literature Review

Research conducted on the long-term positive effects of early interventions by authors such as Campbell et al. (2002) and Heckman et al. (2009) have suggested that early interventions can result in lasting long-term impacts. Campbell et al. (2002) evaluated the long-term outcomes of the benefits of early childhood education intervention for children from disadvantaged areas. This study was a randomised control trial implemented in the United States of America called Abecedarian. Abecedarian was an early-childhood-education intervention implemented in two treatment phases: preschool (the five years from infancy to kindergarten) and school age (the three years in the primary grades). The school-age phase was implemented to improve learning in the primary grades with enhanced parental engagement and tailored support in the classroom. Campbell et al. (2002) conducted a follow-up impact evaluation on Abecedarian on learners who were enrolled in the intervention as young adults (age twenty-one).

The findings from the follow up study showed that participants in the preschool treatment group earned significantly higher scores on intellectual and academic measures as young adults; those in the preschool treatment also attained more years of total education, were most likely to attend a four-year college, and showed a reduction in the rate of teenage pregnancy compared to participants who were part of the preschool control group. Participants from the preschool treatment group showed meaningful effect sizes on reading and mathematics skills, and this impact persisted into adulthood (Campbell et al., 2002). The school-age intervention on its own (learners exposed solely to the school-age intervention and not both school-age and preschool intervention) had a weaker impact than did the preschool intervention on its own (learners ex-

[5] The READ annual is A report produced by the **READ Educational Trust**, a South African non-profit organisation that focuses on improving literacy and reading in schools, particularly in under-resourced areas.

posed only to the preschool intervention). Early-grade interventions aimed at improving early academic skills resulted in increased years of post-secondary education (Campbell et al., 2002).

Heckman et al. (2009) estimated the rate of return of the HighScope Perry Preschool Program, a preschool intervention implemented in Michigan in the early 1960s. The HighScope Program was an early childhood education randomized control trial that aimed to support children's cognitive and socio-emotional development through active learning (Heckman et al., 2009). This program was implemented for three-year-old learners in Perry Elementary school. The intervention lasted for two years and consisted of weekly home visits by teachers as well as a 2.5-hour preschool program in which learners were encouraged to plan, carry out, and reflect on their activities through a "plan-do-review" process. They were also encouraged to make choices, solve problems, and engage in activities. Adults who were part of the treatment were encouraged to observe, support, and extend the children's play appropriately (Heckman et al., 2009).

The results from the follow up study showed an estimated annual social return rate of between 7% and 10%, which was statistically significant for both young men and young women. This rate was well above the historical return to equity rate of ~5.8%. The results also showed that learners who were part of the treatment group had higher long-term educational attainment, increased lifetime earnings, and lower reliance on welfare compared to learners from the control group. These results suggest that early childhood programmes like the Perry Preschool program result in strong, statistically significant social returns (Heckman et al., 2009).

Rea and Burton (2020) critically examined the Heckman Curve and found very little evidence to support some of the claims made by Heckman. Using a large dataset of benefit-cost ratios from the Washington State Institute for Public Policy, the authors showed that interventions that targeted older youth and adults could be just as cost-effective, or even more, than those aimed at young children. They argued that the Heckman Curve was based on a limited and selectively cited set of programmes, lacked the rigor of modern meta-analysis, and oversimplified the complex dynamics of program effectiveness. The study concluded that investment decisions should not be based on the age of recipients alone, but on rigorous, case-by-case evaluations of program impacts and costs.

Bai et al. (2020) conducted a study on the long-term effects (fadeout or growth) of early childhood programmes through eighth grade with evidence from two interventions implemented at scale in North Carolina—The "Smart Start" and "More at Four" programmes. Smart Start was implemented in 1993 and provided state funding to support high-quality early childcare in local communities by allocating money to country-level entities established to improve the quality and availability of early childcare services for learners aged four and younger. The program's purpose was to ensure that all children entered school healthy and ready to learn.

More at Four was implemented in 2001 with the aim of promoting school readiness by providing a high-quality education experience for four-year-old children from disadvantaged areas

in North Carolina. More at Four provided state-funded slots for a year of pre-kindergarten for high-risk four-year-old learners from classrooms. The Smart Start programme spent $52 million in childcare related activities and assisted 1974 childcare programmes to improve the quality of their care: 24,427 learners were provided with subsidies, and 4,732 teachers received an education-based salary supplement (Bai et al., 2020). The More at Four programme spent $5,300 on each kindergarten slot and enrolled 22% of four-year-olds in North Carolina in 2016.

Bai et al. (2020) estimated the long-term impacts of both the Smart Start programme and the More at Four programme by measuring the association between state funding to each programme and student performance after thirteen years of implementation. The results of the evaluation showed significant positive impacts of funding for both programmes on reading and math test scores as well as reductions in grade repetition. The findings also showed that these effects did not "fade out" but rather grew over time as students progressed through middle school. Learners from economically disadvantaged backgrounds showed had larger benefits from the More at Four programme (Bai et al., 2020). In summary, though international studies have been reported in the literature, similar research from the global South is limited.

1.3 Program Description

The Early-Grade Reading Study

In 2015 the Department of Basic Education, Republic of South Africa and various stakeholders implemented and evaluated the Early Grade Reading Study in 230 Quintile 1-3 schools across two districts in North West Province. The program was implemented to improve learner performance in early-grade literacy. The Early Grade Reading Study was a large-scale educational impact evaluation designed to build evidence on successful techniques to improve the teaching and learning of early-grade reading in African languages in the country (Stern et al, 2023). The theory of change of the Early Grade Reading Study suggested that the development of reading comprehension in the Foundation Phase depended upon mastery of vocabulary and decoding and was the gateway to all learning. Although vocabulary can be attained naturally through speaking and hearing, decoding must be taught systematically. Improvement in decoding skills, like the ability to read and understand text, require refinements in the ways education is delivered, including changing the "instructional core," which is the dynamic relationship between the teacher, the student, and the resources they use (books or technology, e.g.). A focus on these three elements can drive meaningful changes in education (Department of Basic Education, 2021).

The Early Grade Reading Study was a randomised control trial composed of three intervention arms and one control arm, aimed at improving early-grade reading in the Setswana Language. The three interventions were implemented with teachers of a cohort of students in Grade

1 in 2015 and continued with the teachers of this cohort in Grade 2 in 2016 and in Grade 3 in 2017 (Department of Basic Education, 2019). The three treatments are outlined below:

Treatment 1: Teacher Training Intervention

This treatment provided teachers with lesson plans aligned with the National Curriculum Statement Grades R-12 as well as the Curriculum and Assessment Policy Statement (Department of Basic Education, 2019). The lesson plans included the use of learning-support materials, which included government-provided workbooks and graded reading booklets, flashcards, posters, etc. The graded readers helped teachers with group-guided reading and individual work. Through this intervention arm, teachers were trained to use the lesson plans and additional learning materials through centralised training sessions held over two days twice yearly.

Treatment 2: On-Site teacher Training and Coaching Intervention

Teachers were given the same set of instructional materials used in Treatment 1 (i.e., scripted lesson plans, graded reading booklets, etc.). However, over and above these materials, teachers also received a one-day training/orientation at the beginning of each term, along with specialist onsite coaching by specialist "reading coaches." In addition to on-site coaching visits, the intervention also offered occasional needs-based workshops.

Treatment 3: Parental Involvement Intervention

Treatment 3 involved weekly parental meetings to discuss the importance of reading in the early grades and to provide parents with the knowledge and tools they needed to become more involved in their children's literacy development. However, this intervention arm was stopped in 2016 after two years of implementation showed little impact. The parents of Grade 3 learners were not part of the intervention and, as a result, this treatment arm was not included in this sustainability research study.

Each intervention was implemented in a group of fifty separate schools, with an additional eighty schools, in which ordinary schooling continued, as controls. After two years of implementation, the training and parental-involvement interventions showed a very small and not statistically significant positive impact. The coaching intervention showed a significant positive impact, and learners who received two years of the intervention were 0.25 standard deviation ahead of control group learners (Department of Basic Education, 2019).

Early Grade Reading Study Learner Sample and Assessments

Table 2 outlines the Early Grade Reading Study learner sample across all five waves of data collection to track students learning outcomes.

In order to increase power and balance among intervention arms, stratified randomization was performed. Ten strata of twenty-three similar schools were created based on school size, socioeconomic status, and performance in the Annual National Assessments. Within each stratum, give schools were then randomly assigned to each intervention group and 8 to the control group, for a total of fifty schools randomly assigned to each intervention and eighty to the control

group. Given that data was collected on twenty Grade 1 learners per school, this sample was sufficient to identify a minimum detectable effect size of 0.21 standard deviation when comparing an intervention group with the control group and a minimum detectable effect size of 0.23 standard deviations when comparing two intervention 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 (Department of Basic Education et al, 2018).

The initial sample of learners (4,538) who were assessed and surveyed in Wave 1 (baseline data collection) were tracked until Grade 4 and Grade 7. The number of learners who were successfully matched and reassessed in the other waves is shown in the Table below. In Wave 5, only 2,424 learners were successfully matched across the other waves.

Table 2: Total Learners Assessed Across All Five Waves

Wave	Number of Learners Assessed	Grade	Data Collection Dates
1	4538	1	2015
2	4137	1	2015
3	3804	2	2016
4	3321	4	2018
5	2424	7	2021

In all five waves of data collection, learner assessments as well as contextual tools for principals, teachers, and parents were completed. Table 3 shows which the specific assessments and questionnaires administered in each wave.

Table 3: Learner Assessments and Questionnaires Administered in Each Wave

Wave	Oral and written Learner assessment	Learner Wellness survey	Teacher question-naire	Teacher Assessment	Principal Question-naire	School Functional-ity survey	Home Back-ground question-naire
1	X		X	X	X		X
2	X		X	X	X		X
3	X		X	X	X		X
4	X				X	X	
5	X	X			X	X	

The insights from this research on the long-term impacts of the Early Grade Reading Study on learner outcomes can allow for the refinement of existing structured pedagogy programmes

implemented across other countries in Sub-Saharan Africa to enhance programmes' long-term effectiveness.

1.4 Empirical Strategy

We adopted a quantitative research approach, and our data were drawn from the Early Grade Reading Study I dataset with a focus on learner data. A longitudinal analysis was conducted on a sample of students that were part of all waves of data collection. Secondary School learner-performance data were retrieved from the national administrative data system known as the Data Driven Districts School Based Assessment dataset. Learners from the intervention were matched with their high school and primary school records, and the impact of the Early Grade Reading Study was tested with the following estimation equation:

$$y_{isgb1} = \beta_0 + \beta_1 Coach_s + X'_{isgb0} + \rho_b + \varepsilon_{isgb1} \quad (1)$$

where y_{isgb1} is the post-treatment outcome indicator of interest for student i in school s in grade g and strata b; Coach is the treatment dummy, which is equal to 1 for the coaching treatment arm and 0 for the control group; X'_{isgb0} is a vector of control variables[6] at baseline, ρ_{b} are strata fixed effects, and ϵ_{isgb1} is the post-treatment error term clustered at the school level (standard errors are clustered at the primary school level because this is the unit of treatment assignment in the randomisation, and outcomes for learners within the same school may be correlated). We used the same estimation equation as did the Stern group (2024). Descriptive statistics on academic performance by grade were also compiled along with correlational analyses of learner performance from primary school grades in the Foundation Phase (Grade 1-3) against outcomes in secondary school grades (Grades 8 and 9).

Evidence of differential attrition was examined in this research as done in the study by Stern et al (2024). Attrition was assessed in two stages. In the first stage, students who were part of the final sample were assessed on whether their characteristics differ from baseline characteristics of the students that were part of the baseline in 2015. The second stage looked into differential attrition by looking at the likeliness of students from the intervention sample dropping out compared to learners from the control group. The implications of attrition were further investigated with a regression analysis on attrition, age, learner gender, and learner baseline reading outcome/score to determine whether the effect of the coaching treatment on learner outcomes varies significantly depending on the level of attrition. Issues such as contamination of the control group as well as the Hawthrone effect will be assessed and discussed in the paper. The implications of

[6] Control variables for Equation 1 included a vector of students' baseline reading scores; student gender, and parental education; school district and average performance on 2014 Annual National Assessments; and average wealth index and attendance rates in the community surrounding the school.

Early Grade Reading Study learners moving out of the province from which the treatment was initiated (North West Province) are also analysed and discussed.

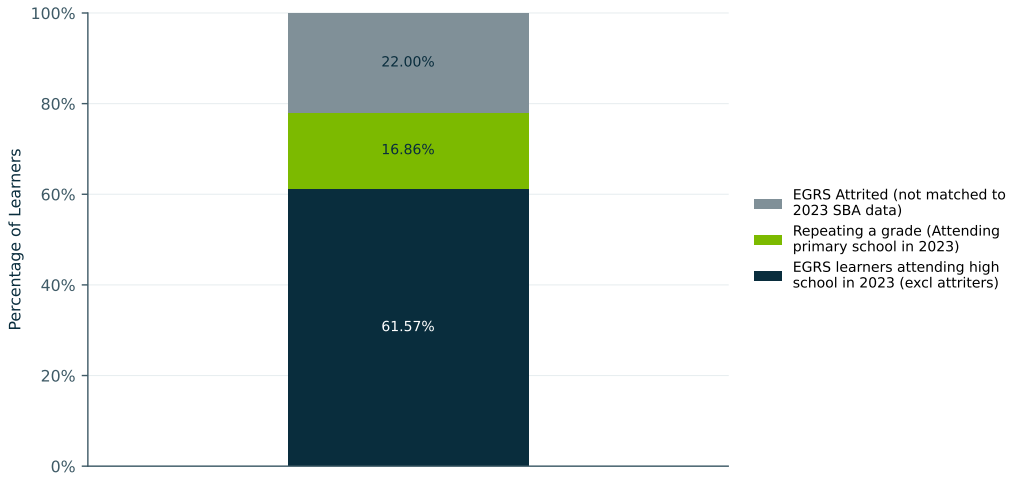
The main regression sample (sample used for estimations) consists of 1,533 learners drawn from the 2,794 learners who were matched to School-Based Assessment data and enrolled in Grades 8 or 9 in 2023. These learners have complete information for all six high school subjects used to construct the composite index, as well as non-missing baseline controls. For the on-track analysis in Table 13, we included 2,024 learners for whom we could determine grade-progression status, following the definition described in the sections to follow.

1.5 Results

Early Grade Reading Study Learner Sample

Under the assumption of uninterrupted grade progression, the Early Grade Reading Study learners would be in Grade 9 in 2023. Figure 2 presents the number of learners successfully matched to the School Based Assessment data, distinguishing between those who were lost to follow-up and those identified as having repeated a grade. The total number of learners assessed in Wave 1 was 4,538; of these, 979 (22%) could not be matched to School Based Assessment data as a result of dropout, a mismatch of ID numbers, or a mismatch of learners' names across the Early Grade Reading Study and School Based Assessment dataset. Seven hundred and sixty-five learners (16.86%) were identified as repeaters and were still in primary school in 2023. The total number of learners who were successfully matched to School Based Assessment data and were in High School In 2023 was 2,794 (61.57%) of the full 4,538 Wave 1 sample. As a result, the long-term impact of the Early Grade Reading Study was estimated using the sample of 2,794 learners in Grades 8 and 9 in 2023.

Figure 2: Distribution of the 2023 Early Grade Reading Study Learner Sample Relative to the Original Wave 1 Cohort



Figures 3, 4, and 5 show the quintile, provincial, and gender distribution of the high- school-learner sample. The majority of learners in the control and treatment arm were in Quintile 1 schools (36% and 39%, respectively). A very small proportion of learners in both the control group and the coaching treatment were in Quintile 4 and 5 schools (9% of control learners and 6% of treatment learners in Quintile 4) and much lower percentages for Quintile 5. From the sample of learners who were identified in School Based Assessment data, some learners were found to have moved to a different province for high school. Most learners who left North West Province moved either to Gauteng Province or to Northern Cape across both control and treatment groups. As shown in Figure 5, the percentage of young women learners was higher in both the treatment and control arm.

Figure 3: Early Grade Reading Study Learners Enrolled in High School in 2023: Distribution by School Quintile and Treatment Arm

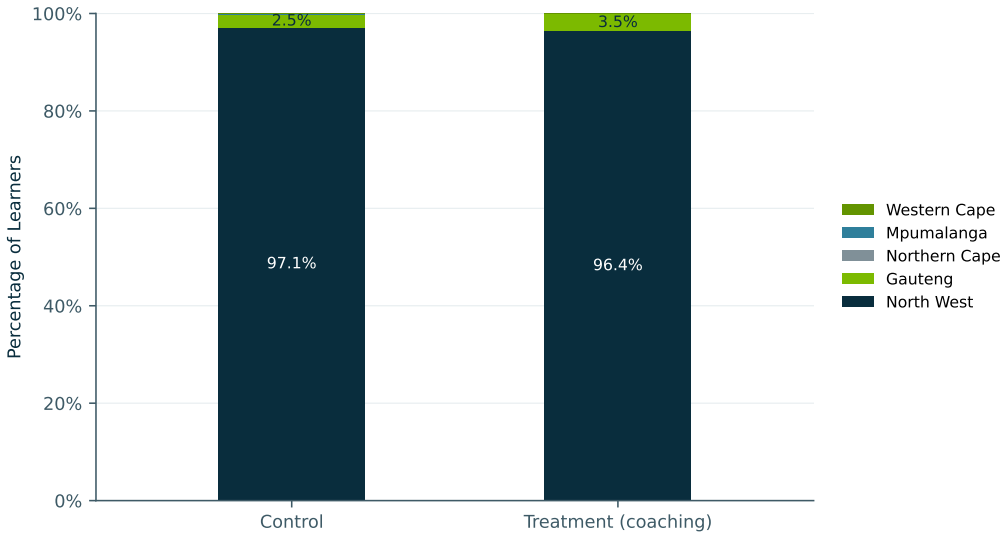


Figure 4: Early Grade Reading Study Learners Enrolled in High School in 2023: Distribution by School Province and Treatment Arm

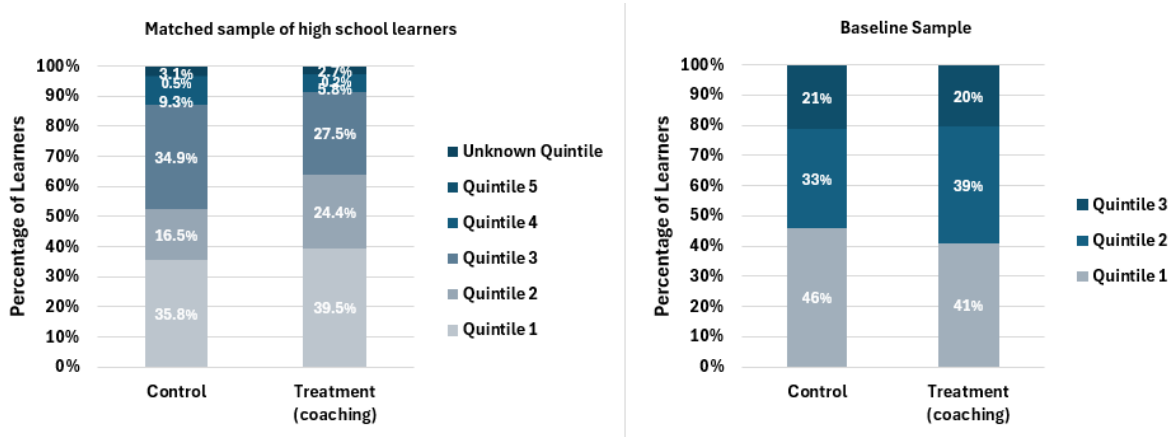
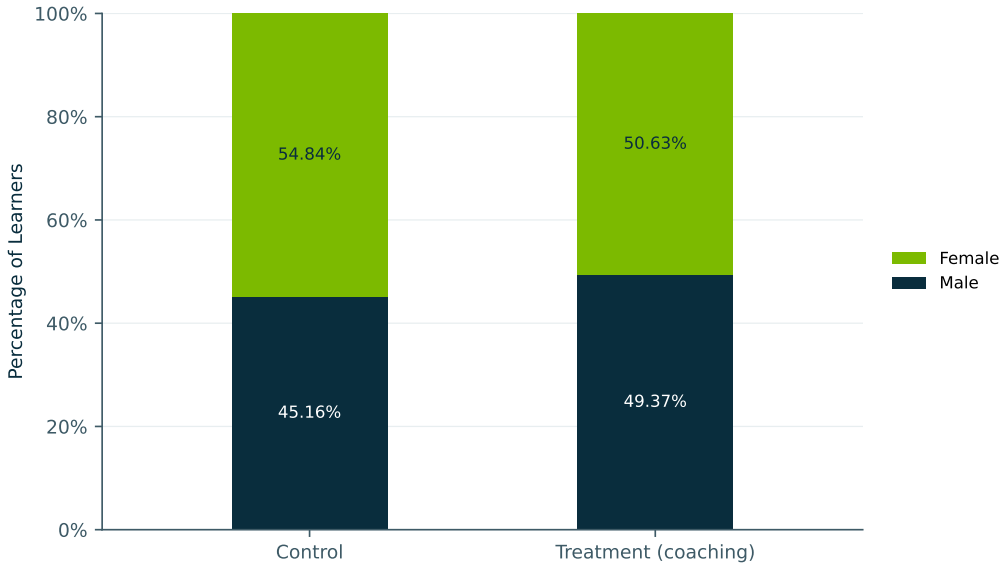


Figure 5: Early Grade Reading Study Learners Enrolled in High School in 2023: Distribution by Gender and Treatment Arm



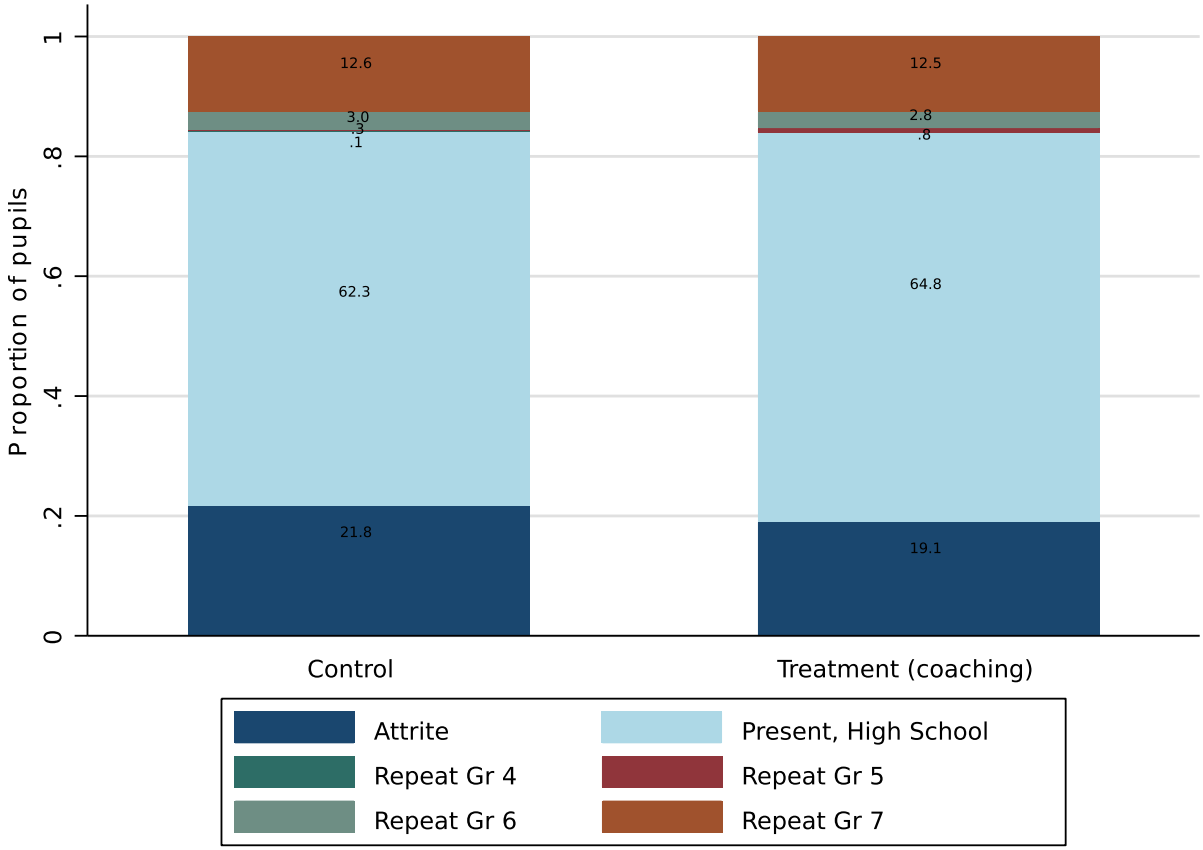
Attrition Analysis and Heterogeneity Tests

Figure 6 shows grade repetition and attrition results from 2023. Attrition rates were at 22% and 19% in the control and treatment groups, respectively. Sixty-five percent of learners from the intervention group were in high school in 2023, compared to 62% of learners in the control group.

As a further step to investigating the effects of attrition, additional analyses and balance tests were conducted, as shown in Table 4. Each column in Table 4 represents a separate regression on attrition, learner age, learner gender, and learner baseline score, respectively. The first Column shows that there was a very small difference in attrition rates between the evaluation arms and the difference is not statistically significant suggesting that attrition is not driven by treatment assignment, but it is systematically related to learner characteristics where lower performing, younger and young men learners are more likely to attrition. Columns 2, 3, and 4 in Table 4 examine whether attrition was selective on observable learner characteristics (age, gender, and baseline learning) and whether this selection differed between treatment and control. The interaction between attrition and the coaching treatment was not statistically significant across all baseline characteristics (and thus not differentially selective across treatment status). This finding suggests that attrition is unlikely to bias estimated treatment effects, however. Figure 6 shows that attrition is slightly higher in the control group, a difference that could slightly bias the results upward.

Additional analyses were undertaken to examine the robustness of the results and explore potential heterogeneity in treatment effects across provinces. Throughout the study, results are

Figure 6: Proportion of Attrition and Grade Repetition by Treatment Group



reported using a simplified composite score[7] that has been standardized with a mean of zero and a standard deviation of one in the control group. As a result, all coefficients from this composite score (High School Composite score) can be interpreted as effect sizes of the intervention. Table 5 below shows a regression designed to test heterogeneity in the treatment effect by province, proxied by whether the learner moved out of North West Province, where the original randomized control trial was conducted, to a different province. Each column represents a separate regression of the treatment effect on composite score (Column 1), English as First Additional Language score (Column 2), and Setswana Home Language score (Column 3) for movers and non-movers.

The regression results show that only 3% of the sample moved out of North West Province and that no statistically significant difference existed in the treatment effect between movers and stayers, shown in the interaction term-Moved out NW x Coaching. Although the interaction term was not statistically significant, the coefficients were quite large, suggesting that potential differences in treatment effects across locations warranted caution in terms of generalizability.

[7] The composite score was constructed from the performance in all the chosen high school subjects (English as First Additional Language, Mathematics, Setswana Home Language, Social Sciences, Natural Sciences, and Economics Management Science) using Principal Components Analysis, a type of factor analysis that identifies the common underlying variation within a set of variables and regards the first principal component in this variation as reflecting the underlying construct, which this case would be learner performance in high school.

Table 4: Attrition Analysis And Balance Tests for Difference in Age, Gender, and Baseline Learning

	(1)	(2)	(3)	(4)
VARIABLES	Attrition	Age	Young women	Baseline Learning
Coaching	-0.0229	-0.00507	-0.0244	0.0901
	(0.0221)	(0.0520)	(0.0225)	(0.162)
Attrition		0.361***	-0.0857***	-0.0968*
		(0.0700)	(0.0244)	(0.0524)
Attrition x Coaching		-0.0511	0.0499	-0.142
		(0.0939)	(0.0437)	(0.138)
Age	0.113***			
	(0.0159)			
Young Women	-0.0333**			
	(0.0129)			
Baseline Learning	-0.0206***			
	(0.00743)			
Observations	2,547	2,547	2,556	2,556
R-squared	0.058	0.048	0.006	0.087
Mean attrition	0.218	0.218	0.218	0.218

*Notes: Each column represents a separate regression, including strata fixed effects. The sample includes all students who were successfully matched to School Based Assessment data. Attrition is a dummy variable equal to 1 if a student was not matched to School Based Assessment data. Standard errors are clustered at the school level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.*

A robustness check was conducted to test whether the treatment effect estimate would change significantly for non-movers compared to movers. The results appear in Table 6.

Table 6 shows the results of the treatment effect separately for learners who stayed in North West Province for high school and learners who moved away. Column 1 represents non-movers and Column 2 represents movers. The regression results show a statistically significant relationship between the treatment effect and the composite scores of both movers and non-movers; this relationship is negative for movers, however, suggesting that the treatment effect may not only weaken but reverse. This result could be due to the small sample size of movers. Table 5, for example, shows that no statistical difference existed in treatment effects between movers and non-movers.

In response to the results from the robustness check, we conducted a balance test to investigate whether treatment and control groups were comparable within movers and non-movers. The results are shown in Table 7. Columns 1-4 represent non-movers and Columns 5-8 include movers and exclude non-movers. The results from Table 7 show that there is no significant difference for most of the dependent variables for both movers and non-movers suggesting that

Table 5: Provincial Heterogeneity Test

VARIABLES	(1) Composite score	(2) Mean English as First Additional Language Score	(3) Mean Setswana Home Language Score
Coaching	0.186***	0.909	1.764*
	(0.0695)	(1.081)	(0.930)
Moved out of North West Province	0.202	-0.855	0.290
	(0.250)	(2.761)	(3.959)
Moved out of North West Province x Coaching	-0.428	1.948	-3.076
	(0.299)	(3.485)	(5.094)
Observations	1,533	1,570	1,562
R-squared	0.152	0.129	0.112
Mean moved	0.0285	0.0285	0.0285

*Notes: Each column represents a separate regression, including strata fixed effects. Control variables are explained in Footnote 7. The dependent variables in Column 1 is the composite score generated for overall performance in high school; the dependent variable in Column 2 is the average score in English as First Additional Language, and the dependent variable for Column 3 is the average score in Setswana Home Language. ***p <.01; **p <.05; *p <.1. Standard errors are clustered at the school level and are reported in parentheses.*

Table 6: Provincial Robustness Test

VARIABLES	(1) Composite Score	(2) Composite Score
Coaching	0.194***	-1.423**
	(0.0697)	(0.689)
Observations	1,493	40
R-squared	0.151	0.709

*Notes: Each column represents a separate regression, including strata fixed effects. Control variables are explained above. The dependent variables in Columns 1 and 2 are the composite score generated for overall performance in high school. Column 1 includes learners who stayed in North West Province in 2023, and Column 2 includes only learners who moved away. ***p <.01; **p <.05; *p <.1. Standard errors are clustered at the school level and are reported in parentheses.*

randomization was preserved when learners moved to a different province. This supports the validity of comparisons between the two groups. Parent education shows a statistically significant relationship; however, this variable is already controlled for in all regressions for this study. This balance test confirms the internal validity of the study.

Table 7: Balance Test for Movers and Non-Movers

	Non-Movers				Movers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Baseline Learning	Young women	Age	Parent Education	Baseline Learning	Young women	Age	Parent Education
Coaching = 1	0.0772	-0.0386	0.0249	0.156**	-0.0540	-0.158	-0.212	0.382
	(0.172)	(0.0273)	(0.0572)	(0.0690)	(0.413)	(0.144)	(0.149)	(0.314)
Observations	1,566	1,566	1,564	808	51	51	fifty	30
R-squared	0.090	0.005	0.009	0.020	0.216	0.102	0.285	0.316
Mean moved	0.0285	0.0285	0.0285	0.0285	0.0285	0.0285	0.0285	0.0285

*Notes: Each column represents a separate regression, including strata fixed effects. Control variables are explained in Footnote 4 above. Columns 1-4 represent non-movers (learners who stayed in North West Province) and Columns 5-8 include movers (Learners who moved to other provinces) and exclude non-movers. ***p <.01; **p <.05; *p <.1. Standard errors are clustered at the school level and are reported in parentheses.*

Table 8 shows the results of an investigation of the external validity of the study by running a regression to test heterogeneity of the treatment across different provinces. North West Province is the base category for this regression. The regression results show that the effect of the treatment varied significantly across provinces. In Gauteng, the result suggests that the effect of coaching for learners who moved to Gauteng is significantly lower than among those who remained in North West Province. The interaction term for the Northern Cape x Coaching is significant and positive, suggesting that the treatment was effective particularly for learners who moved and attended high school in the Northern Cape. The results from Table 9 suggest heterogeneity in treatment effects across provinces. As a result of this, province was controlled for in the overall estimation of treatment effects.

Table 9 shows the regression results of a balance test to investigate whether treatment and control groups were comparable among high-school learners. No statistically significant difference in gender composition, age, or baseline scores between treatment and control groups was noted, suggesting that baseline variables were balanced between the treatment and control groups of the Grade 8 and 9 learners. It can therefore be concluded that randomisation was maintained when the Early Grade Reading Study learners moved to high-school-treatment and that control groups remained comparable. The results of this regression support the validity of using Model 1 to estimate the impact of coaching on high-school learner outcomes.

Impact of Coaching Intervention on Grade 8 and Grade 9 Learner Outcomes

The mean estimates for a sample of high school subjects are shown in Table 10. A graphical representation of these estimates appears in Figure 7. The estimates in Figure 7 demonstrate that

Table 8: Treatment Heterogeneity by Province (Test For External Validity)

VARIABLES	(1) Composite Score
Coaching = 1	0.188***
	(0.0696)
Province-Gauteng	0.230
	(0.250)
Province-Northern Cape	-0.355**
	(0.174)
Coaching x Gauteng	-0.525*
	(0.297)
Coaching x Northern Cape	1.316***
	(0.237)
Observations	1,533
R-squared	0.153

*Notes: Column 1 represents a separate regression of treatment dummies and their interaction with the 2 provinces to which learners moved from North West Province, including strata fixed effects. See Footnote 4 for list of control variable. ***p <.01; **p <.05; *p <.1. Standard errors are clustered at the school level and are reported in parentheses.*

Table 9: Balance Test for Treatment and Control Learners in High School

VARIABLES	(1) Baseline Learning	(2) young women_W2	(3) Age	(4) Parent Education
Coaching	0.0714	-0.0405	0.0184	0.162***
	(0.0525)	(0.0255)	(0.0310)	(0.0573)
Observations	1,617	1,617	1,614	838
R-squared	0.088	0.006	0.008	0.020

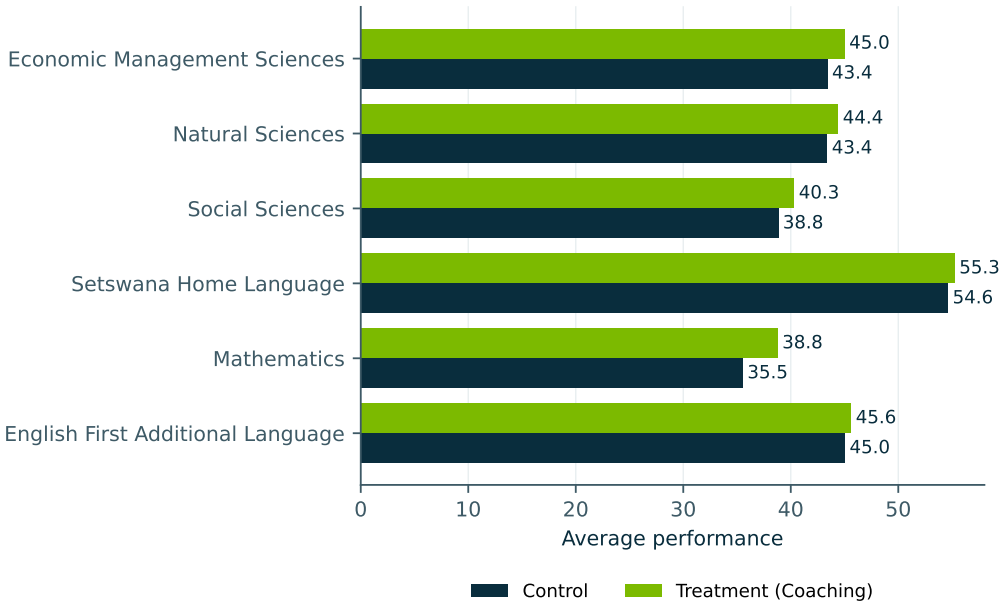
*Notes: Each column represents a separate regression, including strata fixed effects. Control variables are explained in Footnote 4 above. The dependent variables in Columns 1,2 and 3 are baseline learner characteristics namely, baseline scores, gender and Age. ***p <.01; **p <.05; *p <.1. Standard errors are clustered at the school level and are reported in parentheses.*

the average score among learners in the coaching group was generally higher than those in the control group for almost all subjects.

Table 10: Average High School Performance by Subject and Treatment Group

	Control	Treatment (Coaching)	Number of observations
Composite score (standard deviation)	-0.01	0.11	1,533
English as First Additional Language	45.04	45.62	1,570
Mathematics	35.54	38.78	1,609
Setswana Home Language	54.63	55.25	1,562
Social Sciences	38.84	40.32	1,603
Natural Sciences	43.39	44.41	1,603
Economic Management Sciences	43.31	45.01	1,584

Figure 7: Average High School Performance by Subject and Treatment Group



The results shown in Table 10 and Figure 7 suggest that coaching had a small positive impact on overall high school performance. Precision of this treatment effect could be increased through the use of regression analysis to account for factors that differed at baseline. Therefore, the next section presents regression models used to control for factors that might have impacted overall scores but were not affected by the Early Grade Reading Study intervention. All regression models in this section control for the controls listed below along with the addition of province and grade controls.

Table 11 shows an Ordinary Least Squares regression estimate[8] for the impact of the coaching intervention on high-school outcomes, by regressing treatment dummies, controlling for stratification, baseline reading performance, and a set of controls on overall subject performance. Each subject's percentage score was first standardised using the mean and standard deviation of the control group, pooling Grades 8 and 9. We then ran Principal Components Analysis on these six standardised subject scores and retained the first principal component as our composite index of high school performance. Finally, we restandardise this index with a mean of zero and a standard deviation of one within the control group.

Column 2 includes community level controls constructed from the 2022 census.[9] Column 3 includes all other controls with the addition of a control for province. Column 4 includes only Grade 8 learners and Column 5 includes only Grade 9 learners. Column 6 includes all grades and controls with the addition of a control for grade. The motivation for the inclusion of controls was to improve the precision of treatment effect estimates by increasing the explanatory power of the model which lies at 28% for the model in Column 6 from 14% from the model in Column 1. For those learners observed in both Grade 8 and Grade 9 School Based Assessment data, the composite index was computed as the average of the two years; for learners with data in only one grade, we used that year's composite. In our preferred specification, we pooled learners across grades and included a grade dummy to capture average differences between Grade 8 and Grade 9 outcomes. Below is an illustrative representation of the preferred specification (Column 6 of Table 11).

reg composite_score coaching $BL_controls $controls $community_controls i.province i.grade i.strata, cl(primary_School_id)

Table 12 shows Ordinary Least Squares regression estimates (using specifications from Column 6 of Table 11) for the impact of the intervention on overall high-school-subject outcomes, the sample of high school subjects includes Economic Management Sciences, English as First Additional Language, Mathematics, Setswana Home Language, Social Sciences, and Natural Sciences, shown in Columns 2-7, respectively. A graphical representation of this regression can be seen in Figure 8. There, the orange bar shows statistical significance at the 5% level and the green bar shows statistical significance at the 10% level. Table 12 shows that learner's overall performance in high school subjects, as measured by the composite score, increased by 0.12 standard deviation for the treatment group compared to the control group. Columns 2 and 4 show a statistically significant increase in average Economic Management Sciences and Mathematics scores for learners in the treatment group compared to control learners.

[8] Summary statistics tables of the regression samples (reporting means and standard deviations of key variables by treatment status) can be found in Appendix A.

[9] This is a community wealth index derived from several questions in the census about household possessions and the proportion of 13- to 18-year-olds who were attending an educational institution (Department of Basic Education et al., 2018).

Table 11: Ordinary Least Squares Regression Estimates for the Impact of the Intervention on High-school outcomes

VARIABLES	(1) Composite Score	(2) Composite Score	(3) Composite Score	(4) Composite Score	(5) Composite Score	(6) Composite Score
Coaching	0.166**	0.175**	0.175**	0.044	0.174**	0.116*
	(0.072)	(0.068)	(0.068)	(0.077)	(0.073)	(0.061)
Community Controls	No	Yes	Yes	Yes	Yes	Yes
Province Control	No	No	Yes	Yes	Yes	Yes
Exclude Grade 9 learners	No	No	No	Yes	No	No
Exclude Grade 8 learners	No	No	No	No	Yes	No
Control for Grade	No	No	No	No	No	Yes
Observations	1,533	1,533	1,533	677	856	1,533
R-squared	0.142	0.151	0.151	0.126	0.150	0.288
Control mean	1.041	0.947	0.947	0.947	0.434	0.513

*Notes: Each column represents a separate regression of average performance on treatment dummies, controlling for stratification, baseline reading performance, gender, district, Annual National Assessment (ANA) exam performance, and parents' education. Column 2 includes community-level controls constructed from the 2022 census. Column 3 includes all other controls with the addition of a control for province. Column 4 includes only Grade 8 learners, and Column 5 includes only Grade 9 learners. Column 6 includes all grades and controls with the addition of a control for grade. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Table 12: Ordinary Least Squares Estimates for Impact of Intervention on High-School-Subject Outcomes

VARIABLES	(1) Composite Score	(2) Economic Management Sciences	(3) English as First Additional Language	(4) Math	(5) Setswana Home Language	(6) Social Sci	(7) Natural Sci
Coaching	0.116*	2.093*	0.255	2.989**	1.104	1.249	0.641
	(0.061)	(1.237)	(0.951)	(1.352)	(0.887)	(1.282)	(1.324)
Observations	1,533	1,584	1,570	1,609	1,562	1,603	1,603
R-squared	0.288	0.181	0.231	0.137	0.185	0.155	0.187
Control mean	0	43.41	45.04	35.54	54.63	38.84	43.39

*Notes: Each column represents a separate regression estimated using Equation 1. Control variables are explained above. The dependent variables in Columns 2 to 7 are average learner scores in Economic Management Sciences, English as First Additional Language, Mathematics, Setswana Home Language, Social Sciences, and Natural Sciences, respectively. ***p <.01; **p <.05; *p <.1. Standard errors are clustered at the school level and are reported in parentheses.*

Figure 8: Ordinary Least Squares Estimates for Impact of Intervention on High-School-Subject Outcomes

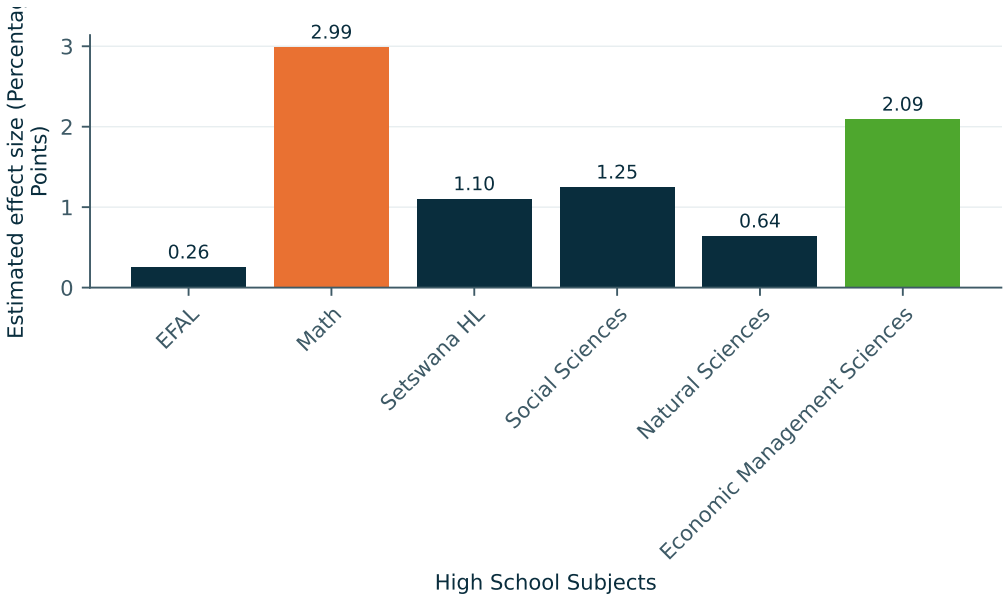


Table 13 explores the link between participation in the intervention and "on track" progress through the first 8 Grades of schooling. Column 1 shows estimates from a linear probability model without any controls and Column 2 includes a full set of baseline, household and community controls. As shown in Column 2, the coaching intervention increases the probability that a learner reaches Grade 9 on time. Specifically, learners who received the coaching treatment are 7.2 percentage points more likely to be on track six years after the intervention than control group learners, a 17% increase relative to the control group.

Table 13: Estimated Effect of the Early Grade Reading Study Coaching Intervention on Student "On-Track" Status

	(1)	(2)
VARIABLES	on_track	on_track
Coaching	0.036	0.072**
	(0.031)	(0.028)
Sample	Non-attriters	Non-attriters
Observations	2,026	2,026
R^2	0.012	0.139
Controls	None	Full
Control mean	0.432	0.432

*Notes: Estimates are from linear probability models. ***p <.01; **p <.05; *p <.1. Standard errors are clustered at the school level and reported in parentheses. The full list of controls can be found in Footnote 4.*

1.6 Discussion and Conclusion

A limited amount of research has been done in low-and middle-income countries on the long-term impact of programmes aimed at improving literacy and numeracy skills in early grades. This study contributed to this research gap by evaluating the long-term impact of an Early Grade Reading intervention on learner outcomes in Grade 8 and 9, using performance in high school subjects such as Mathematics, English as First Additional Language, Setswana Home Language, Social Sciences, Natural Sciences, and Economic and Management Sciences as indicators. The insights into the longevity of the Early Grade Reading Study can allow refinement of existing programmes to enhance their long-term effectiveness, grounding decisions in empirical data rather than theoretical assumptions.

The study examined the long-term impacts of the Early Grade Reading Study approximately six years after the completion of the intervention. The results from the attrition analysis of this study showed evidence of internal validity where the treatment effect was found to be stable regardless of whether or not a learner was lost to follow up. No statistically significant difference was noted in the baseline score of retained participants in the treatment group compared to retained participants in the control group. The findings also showed that a very small sample of Early Grade Reading Study learners (3%) moved out of North West Province. An investigation was conducted to compare learners who moved out of North West Province for high school to learners who stayed in North West Province through a balance test. The results showed that randomisation was preserved when learners moved to a different province, thereby supporting the validity of comparisons between the two groups.

We conducted a further investigation on the external validity of the study by examining whether the treatment effect varied across the provinces that learners moved to compared to its effect in North West Province. The results showed that the effect of the coaching intervention for learners who moved to Gauteng Province was significantly lower than in North West Province. The results also showed that the coaching intervention was more effective for learners who moved and attended high school in the Northern Cape. These differences in the impact of the coaching treatment across provinces may be a result of differences in school quality or instructional alignment and suggests that, although early literacy gains from the Early Grade Reading Study were sustained into high school, they were not sustained equally across provinces. In scaling up this intervention, the alignment of systems and the tailoring of policy responses will be needed for learners who move between education contexts. The small sample size also means that these findings should be interpreted with caution.

Randomisation was maintained as some Early Grade Reading Study learners transitioned to high school, suggesting that the treatment and control groups remained comparable. Because Early Grade Reading Study I was implemented as a stratified randomized controlled trial, schools

were randomly allocated to the coaching treatment within strata, ensuring that treatment and control groups were comparable at baseline.

Our primary estimates focused on learners (specifically those in Grades 8 or 9) who could be matched to School Based Assessment records in 2023, Assuming that, conditional on baseline covariates and strata fixed effects, inclusion in the School Based Assessment sample was not systematically associated with unobserved factors that affected high-school outcomes, the coefficient on the coaching indicator in Equation 1 can be interpreted as the causal intent-to-treat effect for this group of learners.

While the results showed no evidence of differential attrition by treatment on age, gender, and baseline learning, unobserved factors such as parental motivation, health shocks etc. could still influence both attrition and high-school outcomes. These factors could potentially differ by treatment as well.

On average, learners in the coaching group outperformed those in the control group in their high school subjects. Results from the most recent impact evaluation conducted on the Early Grade Reading Study four years after the completion of the intervention showed significant improvement in Setswana Home Language literacy (0.17 standard deviation) in learners in Grade 7 who were exposed to the intervention. The results from this follow-up study conducted six years after the intervention revealed modest but sustained improvements in overall high school performance (about 0.12 standard deviation) and in on-time progression to Grade 9 among learners from coaching schools. Subject-level analysis showed a statistically significant increase in Economic Management Sciences and Mathematics scores for learners in the treatment group compared to control learners. Interestingly there was no impact of the intervention on performance in any of the literacy subjects (English as First Additional Language and Setswana Home Language), suggesting that initial literacy gains in Setswana Home Language observed in 2021, when the learners were in Grade 7, may have faded out by Grades 8 and 9.

Overall, the results from this study support the argument that investing in foundational skills in the early grades can lead to long-term improvements in learner outcomes, particularly through the coaching-intervention approach implemented in the Early Grade Reading Study. Future research could explore the impact of the early-grade intervention on Grade 12 pass rates and on the longer-term educational and economic mobility of learners.

References

Bai, Y., Ladd, H. F., Muschkin, C. G., and & Dodge, K. A. (2020). Long-Term Effects of Early Childhood Programs Through Eighth Grade: Do The Effects Fade Out or Grow? *Children and Youth Services Review*, 112, 104890. https://doi.org/10.1016/j.childyouth.2020.104890.

Bowie, L., Venkat, H., and Askew, M. (2019) Pre-Services Primary Teachers' Mathematical Content Knowledge: An Exploratory Study. *African Journal of Research in Mathematics, Science and Technology Education, 23*(3), 286-297.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., and Miller-Johnson, S. (2002). Early Childhood Education: Young Adult Outcomes from the Abecedarian Project. *Applied Developmental Science, 6*(1), 42-57.

Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, M., and Taylor, S. (2022). The Challenge of Sustaining Effective Teaching: Spillovers, Fade-Out, and the Cost-Effectiveness of Teacher Development Programs. Economics of Education Review, 87(1), 102215. https://doi.org/10.1016/j.econedurev.2021.102215.

Department of Basic Education. (2004). *NATIONAL NORMS AND STANDARDS FOR SCHOOL FUNDING*. https://www.education.gov.za/Portals/0/Documents/Legislation/Call%20for%20Comments/NATIONAL%

Department of Basic Education. (2017). *SACMEQ IV Project in South Africa Report*. https://www.education.gov.za/Portals/0/Documents/Reports/SACMEQ%20IV%20Project%20in%20South%20Africa%20Report.pdf?ver=2017-09-08-152617-090

Department of Basic Education. (2024). *General Household Survey (GHS): Focus on Schooling 2022*. https://www.education.gov.za/Portals/0/Documents/Reports/GHS%202022_Final%2006%2005%2024_2%20Final.pdf?ver=2024-06-07-122826-723

Roberts, N., Moloi, Q. M., & Mort, T. (2024). Assessing student teachers' knowledge of English to inform curriculum design in initial teacher education. *South African Journal of Childhood Education*, *14*(1), a1538. https://doi.org/10.4102/SAJCE.v14i1.1538

Spaull, N., & Taylor, S. (2022). Early Grade Reading and Mathematics Interventions in South Africa-Interventions. *Book*, *3*, 2–24. Cape Town, South Africa: Oxford University Press.

Department of Basic Education, Republic of South Africa (2019). The Early Grade Reading Study Sustainability Evaluation: Technical Report—Improving Early Grade Reading in South Africa. https://www.education.gov.za/Portals/0/Documents/Reports/EGRS/EGRS%20I%20Wave%204%20Report%202019.pdf.

Department of Basic Education, Republic of South Africa (2020a). 2019 School Realities. Department of Basic Education. https://www.education.gov.za/Portals/0/Documents/Reports/School%20Realities%202019%20Final%20.pdf?ver=2020-02-07-101051-330#:~:text=In%202019%2C%20there%20were%2013,see%20Tables%202%20and%203)%20.

Department of Basic Education, Republic of South Africa (2020b). Action Plan to 2024: Towards the Realisation of Schooling 2030. Department of Basic Education. https://www.education.gov.za/Portals/0/Documents/Publications/Action%20Plan%20to%202024%20Brief.pdf?ver=2020-09-14-125337-337.

Department of Basic Education, Republic of South Africa (2021). The Second Early Grade Reading Study Year 4 report: Evidence One Year after the End of Implementation. https://www.education.gov.za/Portals/0/Documents/Reports/EGRS/EGRS%20Indaba%202022/EGRS%20Technical%20Report_.pdf.

Department of Basic Education, Republic of South Africa (2022). The 2022 South African Systemic Evaluation. https://www.education.gov.za/Portals/0/Documents/Reports/2024/SASE%20Info%20Sheet.pdf?ver=2024-12-03-031828-480.

Department of Basic Education, Republic of South Africa, Taylor, S., Mohohlwane, N., Fleisch, B., Prinsloo, C., Van der berg, S., and & Reddy, V. (2016). Early Grade Reading Study: A midline Report. https://www.education.gov.za/Portals/0/Documents/Reports/EGRS/EGRS%20I%20Wave%202%20Report%202016.pdf?ver=2019-05-31-103635-940

Department of Basic Education, Republic of South Africa, Taylor, S., Cilliers, J., Prinsloo, C., Fleisch, B., and Reddy, V. (2018). The Early Grade Reading Study: Impact evaluation after two years of interventions. https://www.education.gov.za/Portals/0/Documents/Reports/EGRS/EGRS%20I%20Wave%203%20Report%202018.pdf?ver=2019-05-31-

Department of Basic Education, Republic of South Africa, Stern, J., Jukes, M., Roper, M., Taimo, L., Bisgard, J., and Williams, B. (2024). Sustainability Impact Evaluation of the Early Grade Reading Study (EGRS) in South Africa. https://www.education.gov.za/Portals/0/Documents/Reports/EGRS/EGRS%20Indaba%202022/EGRS%20

Fleisch, B., Schöer, V., Roberts, G., and Thornton, A. (2016). System-Wide Improvement of Early-Grade Mathematics: New Evidence from the Gauteng Primary Language and Mathematics Strategy. *International Journal of Educational Development*, *49*, 157-174. https://doi.org/10.1016/j.ijedudev.2016.02.006.

Fleisch, B., Taylor, S., Schöer, V., and Mabogoane, T. (2017). Failing to Catch Up in Reading in the Middle Years: The Findings of the Impact Evaluation of the Reading Catch-Up Programme in South Africa. *International Journal of Educational Development*, *53*, 36-47. https://doi.org/10.1016/j.ijedudev.2016.11.008

Fleisch, B. and Alsofrom, K. (2022). Coaching Research in the Early Grade Reading Studies in South Africa. In N. Spaull. and S. Taylor (Eds.), *Early Grade Reading and Mathematics Interventions in South Africa,* 48-64. Cape Town, South Africa: Oxford University Press.

Gove, A., Poole, M. K., and Benjamin, P. (2017). Designing for Scale: Reflections on Rolling Out Reading Improvement in Kenya and Liberia. *New Directions for Child and Adolescent Development Journal. 155*(1), 77-95.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2009). The Rate of Return to the HighScope Perry Preschool Program. *Journal of Public Economics, 94(2010),* 114-128.

Kim, Y. G., Lee, H., and Zuilkowski, S. S. (2020). Impact of Literacy Interventions on Reading Skills in Low- and Middle-Income Countries: A Meta-Analysis. *Child Development*, *91*(2), 638-660. https://doi.org/10.1111/cdev.13204.

Mawoyo, M. and Vally, Z. (2020). Improving Education Outcomes in Low-and Middle-Income Countries: Outcomes-Based Contracting and Early Grade Literacy. *Journal of Learning for Development. 7*(3), 334-348.

Mullis, I., Von Davier, M., Foy, P., Fishbein, B., Reynolds, K., and Wry, E. (2023). *PIRLS 2021 International Results in Reading*. TIMSS and PIRLS International Study Center. https://pirls2021.org/results.https://doi.org/10.6017/lse.tpisc.tr2103.kb5342.

Piper, B., Destefano, J., Kinyanjui, E. M., and Ong'ele, S. (2018). Scaling Up Successfully: Lessons from Kenya's Tusome National Literacy Program. *Journal of Educational Change*, *19*(3), 293-321. https://doi.org/10.1007/s10833-018-9325-4.

Rea, D. and Burton, T. (2020). New Evidence on the Heckman Curve. *Journal of Economic Surveys*, *34*(2), 241--262. https://doi.org/10.1111/joes.12353.

Rodriguez, D. (2022). *Essays on the Promotion of Foundational Literacy and Numeracy in Developing Countries*. Virginia: University of Virginia.

Sparks, R. L., Patton, J., and Murdoch, A. (2013*).* Early Reading Success and Its Relationship to Reading Achievement and Reading Volume: Replication of ' "10 Years Later." *Reading and Writing: An Interdisciplinary Journal, 27*(1):189-211.

Stanovich, K. E. (1986). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, *21*(4), 360-407. https://doi.org/10.1598/RRQ.21.4.1.

Stern, J. M. B., Jukes, M. C. H., Cilliers, J., Fleisch, B., Taylor, S., and Mohohlwane, N. (2024). Persistence and Emergence of Literacy Skills: Long-Term Impacts of an Effective Early Grade Reading Intervention in South Africa. *Journal of Research on Educational Effectiveness, 19*(1), 1-22. https://doi.org/10.1080/19345747.2024.2417288.

Taylor, S. and Spaull, N. (2022). What Works and What Scales? Returning to a Tradition of Evidence-Based System-Wide Programmes. In N. Spaull and S. Taylor (Eds.), *Early Grade Reading and Mathematics Interventions in South Africa* (300-317). Cape Town, South Africa: Oxford University Press.

Taylor, S., Mohohlwane, N., Bohmer, B., and & Murru, A. (2025). Key Findings from a Series of Studies on Structured Learning Programmes in South Africa: Synthesis Report. Department of Basic Education, Republic of South Africa. https://www.wwhge.org/wp-content/uploads/2025/06/WWH-South-Africa-SynthesisReport.pdf.

UNICEF (2022). Improving Early Grade Reading and Numeracy through " 'Catch-Up," a Remedial Learning Programme. https://www.unicef.org/media/124586/file/Zambia%20case%20study%20on%20scaling%20up%20FLNliteracy and numeracy.pdf.

University of the Witwatersrand, and Fleisch, B. (2018). The Education Triple Cocktail: System-Wide Instructional Reform in South Africa. Cape Town, South Africa: UCT Press. https://doi.org/10.58331/UCTPRESS.50fifty

USAID (2015). Status of Early Grade Reading in Sub-Saharan Africa. [Online] USAID. Available: https://www.edu-links.org/sites/default/files/media/file/Status_of_Early_Grade_Reading_SSA_to_USAID.pdf.

Zenex Foundation (2022). Perspectives on Learning Backlogs in South African Schooling. https://www.zenexfoundation.org.za/wp-content/uploads/2022/05/Digital-Perspectives-on-Learning-Backlogs-in-South-African-schooling.pdf.

Appendix A: Additional tables and figures

Table 14: Summary Statistics for Regression Sample

Variable	Control Mean	Coaching Mean	Control Std. Dev.	Coaching Std. Dev.	Control N	Coaching N	Total N
Composite Score	-0.01	0.11	0.98	1.02	929	604	1533
Young women	56%	50%	0.50	0.50	929	604	1533
Mathematics	35%	39%	17.20	18.49	929	604	1533
English as First Additional Language	45%	46%	15.09	15.15	929	604	1533
Setswana Home Language	55%	55%	14.13	14.22	929	604	1533
Social Sciences	39%	40%	16.93	18.16	929	604	1533
Natural Sciences	43%	44%	17.68	18.12	929	604	1533
Economic Management Sciences	43%	45%	17.12	17.30	929	604	1533

Table 15: Balance Baseline Statistics for Regression Sample

Variable	Control Mean	Coaching Mean	Control Std. Dev.	Coaching Std. Dev.	Control N	Coaching N	Total N
Baseline Composite Score	0.14	0.24	1.02	1.14	981	636	1617
Young women	55%	51%	0.50	0.50	981	636	1617
Age	6.39	6.41	0.61	0.59	978	636	1614