LEARNING PROFILES AND READING BENCHMARKS: TRENDS FROM SOUTH AFRICAN DATA ON EARLY GRADE READING

Master of Education 2022

JESS QVIST
Student Number: 1508648

Learning profiles and reading benchmarks:
Trends from South African data on early grade reading

A dissertation submitted in partial fulfilment of the requirements for the degree of
Masters of Education (MEd)
Education Leadership and Policy Studies (ELPS)
University of the Witwatersrand
EDUC7031A

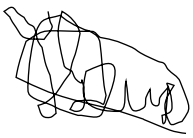
By: Jess Qvist
Student Number: 1508648
March 2023
Supervisors: Brahm Fleisch and Volker Schöer

DECLARATION

I declare that this dissertation, titled “Learning profiles and reading benchmarks: Trends from South African data on early grade reading” is my own work.

It is submitted for the degree of Masters of Education, Education Leadership and Policy Studies, at the University of the Witwatersrand, Johannesburg.

It has not been submitted for any other degree or examination at any other university. All sources have been acknowledged and indicated by means of complete references.



06/03/2023

Jess Qvist

Date

ACKNOWLEDGMENTS

Thank you very much to my supervisors Brahm Fleisch and Volker Schöer for taking the time to guide and encourage me on this journey. You were both always very excited over this topic and my findings and that made me feel like I was on the right track. I also appreciate all your comments as well as the combination of your big picture (Brahm) and small picture (Volker) advice.

Thank you very much to Nic Spaull, Lilli Pretorius and Gabi Wills for answering my emails and providing me with advice on what to research, what dataset to use and how to analyse EGRS data. For a minute there, I had no idea how I was going to contribute to a field where so much incredible progress had already been made and so I really appreciate your guidance.

Thank you very much to Wits for allowing me to complete a Masters of Education in Leadership and Policy (despite the fact that I did not have an undergrad in Education) and for paying for my degree. Education is my true passion and something I used to talk to my friends and family about when I was studying Economics. Back then, I knew that I wanted to learn more about the education system, but I was not sure how to do that, and so I am unbelievably grateful that I received this opportunity.

Thank you to my parents who supported me every step of the way. To my mom for making me countless cups of tea, reassuring me when I was stressed out and in tears, proofreading my paper, guiding me on how to ask for help, listening to me as I explained my research and asking me questions about my research. Thank you to my dad for cooking for me, doing my laundry, asking me questions and for being proud of me. The more I've learnt about the education system, the luckier I feel that I have parents who supported and believed in me from the start. I'd also like to thank you for exposing me to books and teaching me the value of education from a young age. Without that, none of us would be reading this research paper.

Thank you very much to my friend, Kaira Bruggeman for agreeing to be my accountability buddy and for enduring thousands of to-do lists that looked like

- *Wake up at 6*
- *Get up at 6:05*
- *Write a rubbish first draft version, explaining the graphs in the following sections:*
    - *Letter sounds*
    - *Word recognition*
    - *Non-word recognition*
    - *ORF*
    - *Phonological awareness*

It does not need to be perfect! I can edit it later!

Over the last year I've realised the importance of taking small steps and getting started.

Lastly, thank you to all my friends, family and colleagues at Wits who believed in me, asked me how my Masters was going and supported me. I feel blessed that I have such an amazing support system.

DEDICATION

This work is dedicated to all the children in South Africa who are not taught how to read.

Learning how to read was the most empowering, magical and transforming experience in my life and I hope that one day, every child will be given this gift.

ABSTRACT

In South Africa, 90 percent of Setswana learners cannot read for meaning. In this study, I aimed to add to the existing knowledge base by developing learning profiles using the Early Grade Ready Study I longitudinal data. I tracked how Setswana learners with different genders as well as Grade 4 comprehension scores performed in early grade reading components. The findings reveal that learners enter Grade 1 with poor phonological awareness skills, they make slow progress with this skill as well as with letter sounds and word reading in the first grade and greater progress is only made in the second grade. The study also revealed that girls and learners with higher Grade 4 comprehension scores achieved higher phonological awareness, letter sound, word recognition, non-word recognition and oral reading fluency scores in the earlier grades than boys and learners with lower Grade 4 comprehension scores. The implications of this study with regard to teacher pedagogy, assessment and gender equality are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AR	Adult Retrospective
ANA	Annual National Assessment
CCS	Contemporaneous Cross Section
DBE	Department of Basic Education
DIBELS	Dynamic Indicators of Basic Early Literacy Skills
ECD	Early childhood development
EFAL	English as First Additional Language
EGRA	Early Grade Reading Assessment
EGRS	Early Grade Reading Study
FW	Funda Wande
HL	Home Language
k-density	k-density
LCPM	letters correct per minute
LFL	Leadership for Literacy Study
MICS	Multiple Indicator Cluster Survey Programme
NSES	National School Effectiveness Study
ORF	Oral Reading Fluency
PIRLS	Progress in International Reading and Literacy Study
RCT	Randomised Control Trial
RSP	Reading Support Project
SACMEQ	Southern African Consortium for Monitoring Educational Quality
TIMSS	Trends in International Maths and Science Study
TP	True Panel
WCPM	words correct per minute

CHAPTER ONE: INTRODUCTION AND BACKGROUND TO THE STUDY

Literacy is a fundamental part of the right to education. Those who can read have a greater chance of participating in the labour market, earning a higher income, understanding medicine labels and having the ability to fill out forms. In addition, they get to experience the joy and liberation that comes from picking up a book. Despite the multitude of benefits that come from learning to read, UNESCO (2023) estimates that globally, 771 million young people and adults cannot read or write and 250 million children lack basic literacy skills. South Africa is not immune to this crisis. In the 2016 Progress in International Reading and Literacy Study (PIRLS), a nationally representative study that was conducted in 50 countries, it was discovered that 78 percent of Grade 4 learners in South Africa could not read for meaning (Howie et al.,2017). In the 2023 Reading Panel Background report, Spaull (2023) estimates that post-covid, 82 percent of learners cannot read for meaning. The true impact of Covid will only be revealed when the PIRLS 2021 results are released in May 2023.

In the last decade, South African researchers have made great progress in their quest to understand the literacy crisis and the most significant findings are captured in the recently published books, Early Grade Reading in South Africa (Spaull & Pretorius, 2022) and Early Grade Reading and Mathematics Interventions in South Africa (Spaull & Taylor, 2022). While attention was initially given to reading comprehension, in 2019 Spaull and Pretorius introduced the idea of the comprehension iceberg. They explained that comprehension, a skill that was assessed in PIRLS and the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) represented the tip of the iceberg as it was something we could see. It was necessary, however, to dive underneath the surface and explore the bottom of the iceberg. The bottom of the iceberg represents the basic reading components that lead to comprehension. These include skills such as phonological awareness, letter sounds, word recognition, oral reading fluency (ORF) vocabulary and listening comprehension. At the time of their published paper, researchers were only just starting to gather knowledge on these basic reading components, but today they know a lot more. The most notable contributions are the development of letter sound and ORF benchmarks for the Nguni, Sotho-Setswana, English as First Additional Language (EFAL) and Afrikaans languages (Ardington et al 2020, 2021 and 2022; Mohohlwane et al., 2022, Wills et

al., 2022b); the discovery that in Mpumalanga and the Eastern Cape, learners are entering school with underdeveloped phonological awareness and listening skills (Wills et al., 2022a); Setswana and isiXhosa learners are only slowly mastering letter and complex consonant sequences (Wills et al., 2022a); learners across the Nguni and Sotho-Setswana languages have low and slow ORF trajectories, but girls perform better than boys (Wills et al., 2022a); and there is a positive relationship between letter sounds and ORF and ORF and comprehension for Nguni and Sotho-Setswana languages (Pretorius & Spaull, 2022; Wills et al., 2022a).

While significant progress has been made in the reading development of Setswana learners in their early years of schooling, I have identified four gaps in the existing literature. These gaps which are listed below, highlight areas that require further exploration to enhance our understanding of Setswana literacy acquisition and inform targeted educational strategies.

Gaps in the literature:

1. Department of Basic Education Benchmarks: A gap persists in our knowledge regarding the percentage of Setswana learners who attain the Department of Basic Education (DBE) benchmarks for basic reading components throughout their first four years of schooling. This gap underscores the need to understand the trajectory of reading skill acquisition over this critical period.

2. Limited research on phonological awareness, word recognition and non-word recognition: The scant attention given to phonological awareness, word recognition, and non-word recognition for Setswana learners is a gap that necessitates deeper investigation. A comprehensive understanding of these fundamental reading components is vital for tailored instructional strategies.

3. Gender differences in reading performance: The scarcity of research exploring gender differences in word recognition, non-word recognition, letter sounds, and phonological awareness among Setswana learners accentuates the need to discern

potential variations. Addressing gender disparities can contribute to fostering more inclusive and equitable reading instruction.

4. Comprehension and progression relationship: The relationship between comprehension scores in the fourth year of schooling and the pace of progression in earlier years' basic reading components for Setswana learners remains unexplored. This gap prompts inquiry into how early reading development influences later comprehension outcomes.

In light of these gaps, the overarching aim of this research is to contribute substantively to the body of knowledge surrounding Setswana learners' early grade reading development. This aim is pursued through a series of specific research questions that collectively seek to enhance our comprehension of reading proficiency among Setswana learners and the factors that influence it. These research questions are:

1. *After four years of schooling, what percentage of learners who were taught in Setswana have acquired basic reading components (letter sounds, word recognition, ORF)?*

2. *What is the pace of progress in basic reading components (letter sounds, word recognition, non-word recognition, ORF and phonological awareness) in the first four years of schooling for learners who are taught in Setswana?*

3. *Do girls who are taught in Setswana acquire basic reading skills faster than boys who are taught in Setswana?*

4. *What is the relationship between comprehension scores in the fourth year of schooling and the pace of progress in basic reading components in the earlier years for learners who are taught in Setswana?*

This study holds significant importance in the realm of education and literacy development. By addressing the gaps in the current literature surrounding Setswana learners' early grade

reading proficiency, this research aims to offer insights that can revolutionize educational strategies. Understanding the trajectory of reading skill acquisition, exploring gender-based variations, and uncovering the relationship between early reading components and later comprehension outcomes has the potential to not only enhance instructional practices but also contribute to more inclusive and equitable learning environments. The findings of this study can empower educators, policymakers, and researchers with a deeper comprehension of how Setswana learners engage with fundamental reading components, ultimately paving the way for tailored interventions and informed decisions that foster enhanced literacy outcomes for all students.

The conceptual framework that I used to explore the progress in early grade reading components is the learning profile. Learning profiles are graphs that depict the relationship between schooling attainment and learning acquired (Kaffenberger, 2019). In recent decades, many governments have focused their attention on increasing enrolment levels as well as the expansion of schooling. While many more learners have access to schooling, the quality of that schooling is now receiving attention. UNICEF (2022) predicts that globally, the adjusted net primary school attendance rate reached 87 per cent in 2021 with 61 percent of the learners from the poorest quintiles completing primary school compared to 92 of the richest quintiles. Evidence shows, however, that many learners lack basic literacy and numeracy skills. In the sixth round of the Multiple Indicator Cluster Survey Programme (MICS6) which contained a Foundational Learning Skills module that assesses learning outcomes in reading and numeracy skills expected for Grades 2 or 3, findings revealed that only twenty percent of Grade 2 and 3 learners were on-track with literacy and 10 percent with numeracy (UNICEF, 2022). This suggests that the increase in enrolment levels has not necessarily translated into an equal increase in learning. The value of the learning profile is that it measures the amount that children learn in each year of school. Since we can disaggregate these profiles by characteristics such as gender and socio-economic status, we can also determine the groups that are struggling the most (Kaffenberger, 2019). It is highly likely, for example, that children who attend wealthy schools will have steeper learning profiles than those who attend poorer schools. This means that learners in wealthy schools learn more in a year than those in poorer schools.

Kaffenberger (2019) explains that there are three types of learning profiles. The first is Contemporaneous Cross Section (CCS), the second is Adult Retrospective (AR) and the third is True Panel (TP). In the rest of this chapter, I will define these profiles, I will discuss the advantages and disadvantages and I will also provide a few examples of studies that have utilised each of them.

The first learning profile, CCS displays the learning acquired for a cross-section of children across grades at a particular point in time (Kaffenberger, 2019). Researchers adopting this profile, may, for instance, ask learners from Grades 1 to 7 to answer the same numeracy sum and then they will compare the percentage of learners in each grade who got the question correct. If 2 percent of learners got the question correct at the end of Grade 1 and 100 percent of learners got the question correct at the end of Grade 2, the learning profile between the two grades would be steep, indicating that in Grade 2, learners learn how to do this sum. If 3 percent of learners got the question correct at the end of Grade 2, however, the learning profile would be flat, indicating that in Grade 2, learners do not learn how to do this sum. One advantage of using this profile is that it provides information on the current state of the education system (Kaffenberger, 2019). Data is usually also more readily available, it does not take many years to collect the data in the way that panel data does and it may include several grades of schooling. A disadvantage of using this type of profile is that it assumes that different cohorts of learners learn the same amount in each grade. This may not be the case though if one cohort is more academically gifted than another cohort, they are taught by an exceptional teacher or they are following a different curriculum. Examples of researchers who have utilised the CCS profile are Silberstein (2021) and Spaull and Kotze (2015). Silberstein (2021) utilised UNICEF's MICS6 to compare literacy and numeracy scores from the ages 7-14 for 18 countries and Spaull and Kotze (2015) utilised Trends in International Maths and Science Study (TIMSS), the Southern African Consortium for Monitoring Educational Quality (SACMEQ) and the National School Effectiveness Study (NSES) data to assess numeracy performance by socio-economic status in South Africa.

The second profile, AR is a model that captures the retained learning of adults who have acquired various levels of schooling (Kaffenberger, 2019). Researchers using this profile,

may, for example, ask several individuals between the ages of 18 and 50 to complete a literacy and numeracy assessment, and the scores will enable them to measure the health of the education system at various points in time. If the group of participants who completed primary school, but not secondary school twenty years ago cannot read a sentence, it suggests that twenty years ago, learners in primary school did not learn how to read. If they were taught how to read, it suggests that over the next 20 years, they lost this skill. The advantages of this profile are that you can assess retained learning and you can also assess how average scores changed across time (Kaffenberger, 2019). If those who completed primary school but not secondary school ten years ago (as opposed to twenty) could read the sentence, it suggests that between ten and twenty years ago the education system improved. Disadvantages of the AR profile are that it does not provide information on the current education system (unless young individuals are included) and it does not necessarily take into account the possibility that learners may have acquired basic skills outside of the schooling system. Those who did not finish primary school, for example, may have learnt how to read at home or in an adult literacy class. Examples of researchers that used this type of profile are Kaffenberger and Pritchett (2020) who used Financial Inclusion Insights data to analyse the literacy levels of individuals in 10 countries and Pritchett and Sandefur (2020) who used the latest round of the Demographic and Health Surveys data which included a question on literacy for women who had completed less than secondary schooling.

The third profile which can also be referred to as the "pure learning profile" is the TP. This profile utilises panel data which follows the same children across multiple years of schooling (Kaffenberger, 2019). Arguably, this is the most accurate measure of learning as it follows the same children, but since panel data is expensive and takes many years and rounds to obtain results, researchers may not have the option of using this profile (Kaffenberger, 2019). One example of a study that utilised this profile though was conducted by Beatty et al. (2021). These researchers drew upon 2000, 2007 and 2014 data from the Indonesian Life Survey to trace numeracy scores for children between Grade 1 and Grade 12. In this research paper, I used the TP profile because it is the most accurate depiction of learning and South Africa is fortunate to have access to several panel datasets on early grade reading components.

Before moving on to the next chapter, I thought it would be useful to provide some definitions for the key indicators that will be explored in this study.

Alphabetic knowledge – letter sounds and phonological awareness - The concept of alphabetic knowledge is centred on understanding that written symbols correspond to spoken language phonemes (Spaull et al., 2020). This comprehension is pivotal for decoding and early literacy development (Spaull et al., 2020). Alphabetic knowledge is typically assessed through letter-sound relationships, a fundamental skill in early literacy acquisition (Spaull et al., 2020). This knowledge becomes crucial in word reading, enabling children to sound out unfamiliar words based on letter-sound associations (Spaull et al., 2020). It is closely related to phonological awareness, particularly at the phonemic level, which is significant for reading in alphabetic languages (Spaull et al., 2020). Phonological awareness as defined by Smith et al., (2012) is the "ability to recognise the different sound segments of spoken words such as syllables, onset-rimes, and phonemes (p. 12)". Blaiklock (2004, as cited in Spaull et al., 2020) proposes that letter knowledge acts as a mediator between phonological awareness and reading development. As children learn these relationships, they become attuned to individual sounds within words (Spaull et al., 2020). Letter-sound knowledge is a fundamental skill in early literacy development and it takes on a central role in word reading, enabling children to sound out unfamiliar words using their understanding of letter-sound connections (Spaull et al., 2020).

Word and non-word recognition - Reading involves the essential ability to extract meaning from individual printed words, understand their connections within sentences, and construct overall text meaning (Spaull et al., 2020). Proficiency in quickly and accurately reading words forms a cornerstone for successful text comprehension (Spaull et al., 2020). In alphabetic systems, initial letter-sound knowledge is crucial for this process (Spaull et al., 2020). However, true fluency requires recognizing recurring letter patterns, leveraging morphological and orthographic cues, to expedite reading through word chunks (Spaull et al., 2020). Repeated exposure to words leads to familiarity, enabling readers to recognize chunks and develop word-specific knowledge that speeds up reading, freeing cognitive resources for comprehension (Spaull et al., 2020).

Numerous studies establish a strong connection between word reading speed/accuracy and reading comprehension (Spaull et al., 2020). Assessing word reading skills effectively evaluates decoding proficiency (Spaull et al., 2020). Decoding as defined by Smith et al. (2012) refers to "the ability to connect phonemes to letters to sound out unknown words (pg. 12)." Assessments involving context-free word lists and non-words provide insights into phonological recoding ability (Spaull et al., 2020). Non-words lack meaning and orthographic representations, revealing phonological skills without lexical processing (Spaull et al., 2020). Research suggests that real words are processed more swiftly and accurately than non-words, a phenomenon seen across various language structures, including opaque and transparent languages (Spaull et al., 2020).

Oral Reading Fluency (ORF) – ORF encompasses the speed, accuracy, and natural rhythm exhibited by readers when reading aloud, mirroring the intonation of spoken (Spaull et al., 2020). It serves as a broad gauge of reading proficiency (Spaull et al., 2020). While assessing intonation can be complex, oral reading fluency evaluation mainly centres on speed and accuracy (Spaull et al., 2020). Typically, readers read a passage of text in a minute, and errors are deducted from the total words read, yielding a words correct per minute (wcpm) score (Spaull et al., 2020). This is usually followed by an oral comprehension test to assess whether the learner understood the text (Spaull et al., 2020). In South Africa, a strong relationship between ORF and reading comprehension in Setswana has been discovered by researchers (Wills et al., 2022a, 2022b).

The rest of the research paper is structured as follows:

In Chapter Two I review the existing international and South African literature on learning profiles, reading components and the comprehension gender gap. In Chapter Three I outline the methodology used in the study, I describe the EGRS I dataset and the variables that were used in my analysis, I present the descriptive statistics and I discuss the limitations of the study. In Chapter Four, I report the findings of my analysis. In Chapter Five I draw out the main conclusions, I reflect on the limitations of the study, and I discuss opportunities for future research.

CHAPTER TWO: REVIEW OF THE LITERATURE

2.1 Introduction

In this chapter, I will summarise the literature related to the four research questions. As stated in Chapter One, the four research questions that this paper answered were:

> *1. After four years of schooling, what percentage of learners who were taught in Setswana have acquired basic reading components (letter sounds, word recognition, ORF)?*
>
> *2. What is the pace of progress in basic reading components (letter sounds, word recognition, non-word recognition, ORF and phonological awareness) in the first four years of schooling for learners who are taught in Setswana?*
>
> *3. Do girls who are taught in Setswana acquire basic reading skills faster than boys who are taught in Setswana?*
>
> *4. What is the relationship between comprehension scores in the fourth year of schooling and the pace of progress in basic reading components in the earlier years for learners who are taught in Setswana?*

I will start by summarising the international and South African research that has been conducted on early grade reading. Then, I will discuss the literature on the literacy gender gap. Finally, I will explain how I aim to add to the existing literature.

2.2 Reading components

In 2000, the US National Reading Panel published the report, 'Teaching children to read' (Barlett et al., 2015; Sørensen, 2015). The five pillars of reading that were proposed in this report were:

> *1. phonemic awareness, or the ability to identify the individual sounds in spoken words*

2. phonics, or the correspondence of letters (graphemes) to sounds (phonemes)

3. fluency, which is the ability to read text accurately and quickly, with natural prosody

4. vocabulary

5. comprehension, which is the ability to understand and communicate meaning from what is read (Barlett et al., 2015, p. 309; Sørensen, 2015, p. 2)

These five pillars formed the foundation for the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment (Barlett et al., 2015; Sørensen, 2015). Bartlett et al. (2015, p. 309) described DIBELS as a "continuous assessment classroom tool developed for use in the US that reduces reading to discrete skills and then condenses those skills to isolated, quantitative measures… [It] consists of a set of short, timed tests meant to measure phonemic awareness, alphabetic principle, fluency, reading comprehension, and vocabulary."

DIBELS and the five pillars went on to serve as the inspiration behind the Early Grade Reading Assessment (EGRA) which was developed by USAID and RTI International in 2006 (Barlett et al., 2015; Sørensen, 2015). EGRA, as defined by Dubeck and Gove (2015, p. 317) "is a research-based collection of individual subtasks that measure some of the foundational skills needed for reading acquisition in alphabetic languages". The assessment was founded based on the assumption that there are three early stages of reading acquisition (Barlett et al., 2015; Sørensen, 2015). EGRA test components are designed to measure specific stages. The first stage is the emergent stage, and its test components include phonemic awareness, listening comprehension and concepts about print (Gove & Cvelich, 2011). The second stage is the decoding stage, and its test components include letter naming, letter sounds, syllable naming, nonsense word reading and familiar word reading (Gove & Cvelich, 2011). The third stage is the confirmation and fluency stage, and its test components are paragraph reading (ORF) with comprehension as well as dictation (Gove & Cvelich, 2011). The rate at which children pass through these stages is dependent on country and language (Barlett et al., 2015; Sørensen, 2015). While EGRA has been adapted in more than 65 countries and 100 languages, it has received its share of criticism (Dubeck & Gove, 2015). The first criticism is from those who believe that reading is not taught in stages. They hold that comprehension, fluency and decoding skills should be taught simultaneously (Barlett et al., 2015; Sørensen,

2015). Dubeck and Gove (2015) argue, however, that EGRA measures numerous skills that are shown to be consistently strong indicators of comprehension. The second criticism is that EGRA ignores interactions between oral language development (which children are usually exposed to at home through rhymes and songs), print exposure and reading and it also ignores the links between reading and writing (Barlett et al., 2015; Sørensen, 2015). Dubeck and Gove (2015) contend though that EGRA is not intended to measure background characteristics and literacy. A third criticism is that the timing of tasks undermines the demonstration of comprehension of reading (Barlett et al., 2015; Sørensen, 2015). A slow reader may understand what they are reading if they were granted more than a minute, but the EGRA test only allows a minute. One solution to this is providing learners with an extra two minutes to answer comprehension questions (Wills 2022a). A fourth criticism is that EGRA was developed with the monolingual English child in mind and its reading assumptions may not hold in multilingual and other language contexts (Barlett et al., 2015; Sørensen, 2015). This has been attested, however, with the assertion that EGRA is flexible and adaptable, and can be tailored to any language and orthography (Sørensen, 2015).

In South Africa, early grade reading assessments initially focused on reading comprehension. The two most popular studies were PIRLS which is an international assessment designed to measure the reading achievement of Grade 4s (NCES, 2016) and SACMEQ which tests Grade 6 reading achievement in member countries from the Southern and Eastern Africa region (GEM, 2015). The PIRLS results for South Africa indicate that there has been a steady improvement, albeit from a low base (Van Staden & Gustafsson, 2022). In PIRLS 2016, for example, it was found that 78 percent of learners did not reach the low international benchmark. This means that the majority of learners could not:

> *[when reading predominantly simpler literary tests], locate and retrieve explicitly stated information, actions, or ideas, make straightforward inferences about events and reasons for actions or begin to interpret story events and central ideas … [when reading predominately simpler informational tests locate and reproduce explicitly stated information from text and other formats (e.g., charts, diagrams) and begin to make straightforward inferences about explanations, actions, and descriptions*
> (Mullis et al., 2017, p. 60)

In other words, 78 percent of learners cannot read for meaning. While it is valuable to know that the majority of learners in South Africa cannot read for meaning, in 2019, Spaull and Pretorius called for more research on basic reading components. They explained this using the metaphor of an iceberg. The tip of the iceberg (the part that we can see) represents reading comprehension and the bottom of the iceberg (the part below the surface of the ocean) represents basic reading components. These components can be separated into oral language comprehension (the ability to use and understand spoken language) as well as decoding skills ("the ability to accurately read familiar words and decode unfamiliar words out of context" (Spaull & Pretorius, 2019, p. 158)). Skills such as listening comprehension, vocabulary and morphosyntactic knowledge (knowledge of grammar) fall under oral language comprehension, and phonetic awareness (ability to analyse the sound structure of oral language (Korpipää et al., 2017), letter-sound knowledge, word reading and reading fluency fall under decoding (Spaull & Pretorius, 2019). Many of these skills are tested in EGRA.

Prior to this, a few studies had been conducted on various early grade reading components. Two studies that tested vocabulary in the Foundation Phase were Malda et al. (2014) and Wilsenach (2015). Malda et al. (2014) found that in a sample of third graders, vocabulary made a more significant contribution to reading comprehension for the 127 English learners than the 109 Setswana and 122 Afrikaans learners. Wilsenach (2015) observed that in term 3, vocabulary levels were low for the 99 Grade 1 Northern Sotho (also known as Sesotho) learners who either received their instruction in English or Northern Sotho. These two research groups also explored phonological awareness. Malda et al. (2014), noticed that phonological awareness (through the form of phoneme blending, phoneme segmentation, and phoneme deletion skills) played a greater role for the Setswana and Afrikaans learners than the English learners and Wilsenach (2015) found that the development of letter knowledge and phoneme-grapheme correspondences (a phonological awareness task) was on-track for the Northern Sotho learners who had received their instruction in English but it was off-track for those who received instruction in their Home Language.
Another three studies exploring phonological awareness in South Africa were Hugo et al. (2005), De Sousa et al. (2011) and De Witt and Lessing (2016). Hugo et al. (2005) discovered a significant relationship between preschoolers' phonological awareness and later reading

success in Afrikaans Home Language (HL) for 71 learners in three schools. De Sousa et al. (2011) noted a moderate positive relationship between Zulu phonological awareness and English spelling for 30 Grade 2 bilingual Zulu-English children. De Witt and Lessing (2016) measured the impact of the "Growing-to-Read" programme on the phonological awareness skills of Xitsonga preschool children in rural areas. The school readiness programme was shown to have a positive impact on phonological awareness skills for the 223 learners who received the intervention. The control included 70 additional learners.

Researchers in South Africa had also explored the relationship between various reading components. Wilsenach (2016) found, for example, that there were strong correlations with a non-word repetition task administered to 120 Grade 3 Northern-Sotho learners and other measures of phonological working memory, and they reported that performance on the task was associated with word and fluent reading. O'Carroll (2011) noted significant relationships between early Grade 1 letter knowledge and the end of Grade 1 word reading and spelling in two English medium-schools and Schaefer and Kotzé (2019) observed significant relationships between first language phonological awareness and letter sound knowledge in HL for later wording abilities in isiZulu, siSwati and English for 1347 learners.

Despite these important links, Schaefer and Kotzé (2019) reported that siSwati, isiZulu and English learners enter school with varying levels of oral language proficiency and O'Carroll (2011) observed that in two English-medium schools almost half of the learners entering Grade 1 were unable to recognise letter sounds. While Schaefer and Kotzé (2019) noted that there was some improvement in decoding skills, 46 percent of learners were unable to read a single word in their HL in a minute and 35 percent were unable to read a single word in English.

From about the time that Spaull and Pretorius (2019) published their chapter, large-scale research on decoding skills proliferated in South Africa. Large datasets (approximately 40 thousand learners) were created for different language groups by collating several EGRA-type studies (Wills et al., 2022a). These panel data studies included EGRS I in the North-West and EGRS II in Mpumalanga, the Reading Support Project (RSP) in a sub-sample of EGRS I schools, the Story Powered Schools (SPS) project in KwaZulu-Natal and the Eastern

Cape, the Funda Wande (FW) interventions in the Eastern Cape and Limpopo and the Leadership for Literacy Study (LFL) in KwaZulu-Natal, Gauteng and Limpopo (Wills et al., 2022a). Researchers along with the DBE have used these datasets to analyse the relationship between various reading components and to develop benchmarks and thresholds in Nguni languages, (Ardington et al 2020 and 2021) Sesotho-Setswana languages (Mohohlwane et al., 2022, Wills et al., 2022b), English as a First Additional Language (EFAL) (Wills et al., 2022b) and Afrikaans (Ardington et al., 2022). For learners to comprehend text, it is recommended that they reach these benchmarks. Mohohlwane et al. (2022) explained that languages must have their own benchmarks because of differences in phonology, morphology and orthography. The DBE recommends that all Sesotho-Setswana learners should be able to sound at least 40 letters correct per minute (LCPM) at the end of Grade 1 and read 40 and 60 words correct per minute (WCPM) at the end of Grades 2 and 3 respectively for them to comprehend text (Mohohlwane et al., 2022, Wills et al., 2022b). These figures are 40, 20 and 35 for Nguni learners (Ardington et al 2020 and 2021; Mohohlwane et al., 2022) and 40, 50 and 80 for Afrikaans learners (Ardington et al 2022). For EFAL, the DBE recommends that at the end of Grade 2, learners should be able to read 30 WCPM, at the end of Grade 3 they should be able to read 50 WCPM, at the end of Grade 4, they should be able to read 70 WCPM, at the end of Grade 5 they should be able to read 90 WCPM and by the end of Grade 6 they should be able to read 100 WCPM (Wills et al., 2022b). The percentages who are reaching these benchmarks are small. In their paper, Mohohlwane et al. (2022) estimated that in Nguni HL, 7-32 percent of non-representative grade pre-covid samples reached the Grade 1 benchmark at the end of Grade 1 or the start of Grade 2, 29 to 54 percent reached the Grade 2 benchmark at the end of Grade 2 or the start of Grade 3 and 11 to 47 percent reached the Grade 3 benchmark at the end of Grade 3 or the start of Grade 4. They estimate that in Sesotho-Setswana HL, 24 percent of non-representative grade pre-covid samples reached the Grade 1 benchmark at the end of Grade 1 or the start of Grade 2, 36 to 56 percent reached the Grade 2 benchmark at the end of Grade 2 or the start of Grade 3 and 24 to 48 percent reached the Grade 3 benchmark at the end of Grade 3 or the start of Grade 4.

Wills et al. (2022a) also used one of these large datasets to explore phonemic awareness for new Grade 1s in Mpumalanga (isiZulu and siSwati) and the Eastern Cape (isiXhosa), letter-sound knowledge for Grade 1s and 2s in isiXhosa and Setswana and ORF trajectories across

several grades, terms and language groups. They discovered that learners are entering school with underdeveloped phonological awareness skills, they are only slowly mastering letter and complex consonant sequences, and their fluency levels are low. (Over 55 percent of learners cannot read a single word correctly by the end of Grade 1). They also discovered that there was a positive relationship between phonological awareness and ORF for isiXhosa learners as those who got all three phonological awareness questions correct in Grade 1 scored 15 points higher in ORF in Grade 3 than those who got none of the questions correct. Finally, they found a positive relationship between ORF and comprehension. Pretorius and Spaull (2022) built on Wills et al. (2022a) research by documenting Setswana ORF trajectories for three groups of letter decoders. They found that Letter Decoders (those who could sound over 26 LCPM in Grade 1) could read more words across grades than the other two groups, Delayed Decoders (10-25 LCPM) could not read a single word by the end of Grade 1 but they could read 23 WCPM by the end of Grade 2 and Alphabetically Illiterate learners (less than 10 LCPM) could not read a single word by the end of Grade 2 and only reached the Grade 3 benchmark (60 WCPM) in Grade 7.

Research on oral language comprehension remains low although in their study, Wills et al. (2022a) utilised Grade 1 listening comprehension data from FW-EC and EGRS I and found that Eastern Cape (isiXhosa) learners scored an average of 53 percent and Mpumalanga (isiZulu and siSwati) learners scored 55 percent. This suggests that learners are entering Grade 1 with poor listening comprehension skills.

2.3 Gender gap in literacy

Research consistently shows that girls are better at reading comprehension than boys. This trend holds internationally, in Africa and in South Africa. In PIRLS and SACMEQ, for example, girls outperformed boys in the majority of the countries, including South Africa. In PIRLS 2006, Grade 4 girls had significantly higher reading results than boys in every country (there was a total of 40) except Luxembourg and Spain where reading scores were equal (Mullis et al., 2007). In 2011, this trend held, and girls outperformed boys in 40 out of a total of 45 countries. Colombia, Italy, France, Spain and Israel showed no difference (Mullis et al., 2012). In 2016, girls outperformed boys in all 50 countries except Macao SAR and Portugal

(Mullis et al., 2012). In SACMEQ II (2000), the countries with reading scores that were significantly higher for Grade 6 girls than boys were Botswana, Mauritius, Seychelles and South Africa (Saito, 2011). In SACMEQ III (2007), this list expanded to Botswana, Mauritius, Namibia, Seychelles, South Africa and Zanzibar. Contrastingly, boys outperformed girls in Tanzania (mainland) in SACMEQ II and III and Malawi in SACMEQ III (Saito, 2011).

In South Africa, girls outperform boys in literacy, but they have also been shown to outperform boys in every subject throughout school. Spaull and Makaluza (2019), for example, assessed PIRLS, Trends in International Mathematics and Science (TIMSS), SACMEQ and matriculation results between 1995 and 2018 and found that girls outperformed boys at the mean in all subjects. They found that Grade 4 girls are an entire year ahead of Grade 4 boys in reading outcomes, matric girls performed better than boys in all 13 subjects and for every 100 girls in matric in 2018, there were only 80 boys. The reason a smaller proportion of boys writes matric could be explained by the high percentage of boys who are held back during school. Using a Birth-to-Twenty child cohort study that followed children born between April to June 1990 in the Greater Johannesburg area for two decades, Fleisch and Shindler (2009) found that there were high rates of repetition for boys living in working-class and poor neighbourhoods.

In 2022, Hoymeyr published a paper, exploring the reasons girls in South Africa outperform boys in reading and maths. She found that girls progress through the early grades faster than boys and they also achieve higher scores in reading self-efficacy and enjoyment and engagement in reading lessons. This finding is supported by Govender (2020) who interviewed thirteen foundation phase teachers from five English medium primary schools in Port Shepstone, KwaZulu-Natal. The teachers reported that girls are better readers than boys because they are more motivated and confident. Interestingly, Hofmeyr (2022) found that girls did not hold higher levels of motivation in maths.

The question of whether girls outperform boys in basic reading components has not received as much attention. There are, however, a few international and local studies. In Norway, Sigmundsson et al. (2017) tested whether there was a gender gap in letter-sound knowledge when 485 children aged 5-6 years started school and they found that girls were

significantly better at identifying the sound of both capital and lowercase letters. In the United States, Below et al. (2010) used a cross-sectional design to test the difference in gender performance in letter sounds, letter naming, phoneme segmentation fluency, nonsense word fluency and ORF. 1218 children between kindergarten and Grade 5 in three elementary schools in a rural south-eastern school district participated in the study. The researchers found that in kindergarten, girls performed significantly better than boys in the letter sound, phoneme segmentation fluency and nonsense word tasks. In Grade 1, girls were better at the letter naming task but not the phoneme segmentation task. Girls also slightly outperformed boys in the nonsense word task but the difference was not significant. In the ORF task, there were no significant differences between the genders in Grades 1, 2, 3 and 5. In Grade 4, however, ORF for the girls was significantly higher than what it was for the boys although the difference was small.

In South Africa, this question has received attention from two groups of researchers. Wilsenach and Makaure (2018) tested 60 Northern Sotho-English learners in Grade 3 at a primary school with English as their medium of instruction that was located in a high-poverty suburb of Tshwane. The tasks that were administered in English were word recognition, digit span, regular word reading, exception word reading, non-word reading and fluent reading. The researchers found that there was a significant pro-girl gap for every task. The second group, Wills (2022a) examined whether the Setswana fluency trajectory gender gap widened across primary school. They found that at the median point, a widening pro-girl fluency gap was evident into the second year of school, but this gap remained relatively constant until the end of Grade 7. At the 25th and 75th percentiles, however, the pro-girl fluency gap widened across all grades. This means that the gender gap between the top and bottom performers widened as learners progressed through primary school.

2.4 Analysis

In this chapter, I provided an overview of the early grade reading and the literacy gender gap international and local literature. In the first section, I provided a summary of the research that has been conducted on early grade reading both internationally and in South Africa. The EGRA which is founded on the assumption that reading is acquired in three

stages is arguably the most dominant voice in the space. While the EGRA has been adapted in many countries, one of main criticisms is from those who assert that reading is not acquired in stages and that comprehension, fluency and decoding skills should be taught simultaneously. Dubeck and Gove (2015) argue, however, that EGRA is valuable as it measures numerous skills that are shown to be consistently strong indicators of comprehension. It is also flexible and can be tailored to any language and orthography (Sørensen, 2015). The EGRA model has since been adapted in South Africa, and as mentioned, we have seen a rise in the use of panel data testing numerous EGRA reading components. Prior to this, South African researchers only had access to large-scale data on reading comprehension from studies such as PIRLS and SACMEQ. Spaull and Pretorius (2019) explained that reading comprehension is at the tip of the iceberg and there is a greater need to dive underneath the surface of the ocean and explore the bottom of the iceberg: basic reading components that lead to comprehension. Up to that point, researchers had conducted research on basic reading components in South Africa, but the sample sizes were relatively small. Since then, large scale research has been conducted and the DBE has developed letter sound and ORF reading benchmarks for the Nguni, Sotho-Setswana, EFAL and Afrikaans languages. Learners who reach these benchmarks at their specified points stand a greater chance of reading with comprehension in Grade 4. Wills et al. (2022s) also found that learners in Mpumalanga (isiZulu and siSwati) and the Eastern Cape (isiXhosa) are entering school with underdeveloped phonological awareness and listening skills, Setswana and isiXhosa learners are only slowly mastering letter and complex consonant sequences and learners across many grades and African languages have low and slow fluency trajectories. Wills et al. (2022a) and Pretorius and Spaull (2022) have also found positive relationships between letter sounds and ORF and ORF and comprehension in Setswana. Moreover, Schaefer and Kotzé (2019) found significant relationships between first language phonological awareness and letter sound knowledge in HL for later wording abilities in isiZulu, siSwati and English.

It is encouraging that more research has been conducted on basic reading components in South Africa, however, I have identified three gaps in the literature. The first gap is the percentage of Setswana learners reaching the ORF and letter sound benchmarks in the first four years of schooling. While Mohohwlane et al. (2022) calculated the percentage of

learners reaching the benchmarks, they did not explore how this changed across four years of schooling. (I.e. they calculated the percentage of learners that reached the Grade 2 benchmark in Grade 2, but not the percentage that reached it in Grades 1 and 4). The second gap that I have identified is the progression in phonological awareness, word recognition and non-word recognition for Setswana learners in the first four years of schooling. The third gap is the relationship between Grade 4 reading comprehension scores and the pace of progression in letter sounds, word recognition, non-word recognition and phonological awareness in the first four years of schooling for learners who are taught in Setswana. This research paper aims to fill these gaps.

In the second section, I explored the gender gap in reading acquisition. It is widely known that girls have higher reading comprehension levels than boys. This trend holds in numerous countries including South Africa. It is less widely known, however, if this gap holds at the basic reading component level. In Norway, Sigmundsson et al. (2017) found that in a group of learners aged 5-6, the girls performed better than the boys in the identification of capital and lowercase letter sounds. In the United States, Below et al. (2010) found that in three elementary schools in a rural southeastern school district girls in kindergarten performed significantly better than boys in letter sound, phoneme segmentation fluency and nonsense word tasks. They also found that in Grade 1, girls were better at the letter naming task but not the phoneme segmentation task and girls slightly outperformed boys in the nonsense word task, but the difference was not significant. In the ORF task, there were no significant differences in Grades 1, 2, 3 and 5. In Grade 4, however, ORF for girls was significantly higher than for boys although the difference was small. In South Africa, Wilsenach and Makaure (2018) found a significant pro-girl gap in eight tasks that were administered in English (word recognition, digit span, regular word reading, exception word reading, non-word reading and fluent reading) to Grade 3 Northern Sotho-English learners. Wills (2022a) also found that at the median point, a widening pro-girl fluency gap was evident into the second year of school for Setswana learners, but this gap remained relatively constant until the end of Grade 7. At the 25th and 75th percentiles, however, the pro-girl fluency gap widened across all grades. This means that the gender gap between the top and bottom performers widened as learners progressed through primary school. The question of whether this gap exists and changes in the first four years of schooling in phonological

awareness, word recognition, non-word recognition and letter sound tasks has not been answered. I aim to answer this for Setswana learners in my study.

Therefore, my study will make use of learning profiles and EGRS I panel data to explore the percentage of learners who are taught in Setswana and acquire basic reading skills (letter sounds, word recognition, non-word recognition, ORF and phonological awareness) in the first four years of schooling. It will also assess whether these learning profiles differ for each gender. Finally, it will explore the relationship between Grade 4 comprehension scores and reading components in the first four years of schooling.

CHAPTER THREE: RESEARCH DESIGN AND METHODOLOGY

3.1 Introduction

This study took a quantitative approach, and I analysed secondary data. The advantage of a quantitative approach is that the sample size is bigger, and results can be generalised across a larger group of people (Rahman, 2020). A disadvantage, however, lies in the fact that the quantitative method often prioritizes breadth over depth (Rahman, 2020). While it allows us to draw robust statistical conclusions and make generalizations, it may not capture the intricate nuances and qualitative aspects of learners' experiences that a more qualitative approach could offer. Nevertheless, the decision to adopt a quantitative approach is grounded in its unique strengths that align with the objectives of this study. In this section, I will start by describing my data source and the variables that I used in the analysis. Then I will describe the methods that were employed. Finally, I will discuss the limitations of EGRS I data.

3.1 Data source: EGRS I

The dataset that I used in this study was the Early Grade Reading Study I (EGRS I). EGRS I was a Randomised Control Trial (RCT) that was initiated by the DBE in 2015 in the North West (Taylor et al., 2017). The study aimed to assess the effects of three interventions (lesson plans; lesson plans and on-site coaching; and weekly parental meetings on the importance of reading) on various reading components (Taylor et al., 2017). 230 quintile 1-3 schools participated in the trial – 50 in each intervention and 80 formed the control (Taylor et al., 2017). The meticulous selection process of these 230 schools began with an initial pool of 458 primary schools located in the Dr Kenneth Kaunda and Ngaka Modiri Molema districts, as identified from the 2014 administrative data (Taylor et al., 2017). To ensure the representativeness of the sample, certain exclusion criteria were applied. Schools falling under quintiles 4 and 5 were omitted, as were those employing languages other than Setswana for Foundation Phase instruction (Taylor et al., 2017). Additionally, schools absent from the 2014 Annual National Assessment (ANA) dataset were removed, as were eight schools previously selected for piloting project instruments (Taylor et al., 2017).

Furthermore, schools with fewer than 20 Grade 1 learners or more than 180 Grade 1 learners were excluded from consideration (Taylor et al., 2017). Further refinement followed, whereby three additional schools were excluded after review by the North West PED revealed specific issues—such as school closures or management conflicts (Taylor et al., 2017). Following this thorough screening, a total of 235 schools remained eligible for the trial. To ensure a well-balanced sample, five schools were randomly excluded from this pool, ultimately resulting in the final selection of 230 schools for the EGRS I trial (Taylor et al., 2017).

To enhance statistical power and ensure equitable representation across intervention arms, they implemented a method of stratified randomisation, a crucial aspect of the study design (Taylor et al., 2017). Employing this approach, they organised the schools into 10 distinct strata, each comprising 23 schools that demonstrated similarity in terms of school size, socioeconomic status, and performance levels as observed in the ANAs (Taylor et al., 2017). Within every stratum, a rigorous process of random allocation was carried out. They distributed 5 schools randomly to each of the three intervention groups, while the control group encompassed 8 schools (Taylor et al., 2017). This meticulous allocation strategy culminated in the assignment of 50 schools to each intervention arm and 80 schools to the control group, collectively forming the cohort for the study (Taylor et al., 2017).

The rationale behind this careful selection and allocation process was rooted in the need to yield meaningful outcomes. With data collection centered on 20 Grade 1 learners per school, the sample size was calculated to provide the study with the capability to identify a minimum detectable effect size of 0.21 standard deviations when contrasting an intervention group against the control group (Taylor et al., 2017). Furthermore, it allowed for the identification of a minimum detectable effect size of 0.23 standard deviations when comparing two distinct intervention groups (Taylor et al., 2017). Importantly, these calculations hinged on several critical assumptions, including a confidence interval of 95%, an alpha value of 0.8, an intra-class correlation coefficient (rho) of 0.3, and a correlation coefficient between pre- and post-test scores of 0.7 (Taylor et al., 2017). These factors collectively underscore the methodological rigour employed in the study's design, aiming to provide robust and reliable results.

Between 2015 and 2020, five waves of data collection took place (see Table 1 below for details) through the form of teacher and principal questionnaires as well as learner assessments. Every year, starting with a baseline, the same learners were tested in EGRA-type tasks which included letter sounds, word and non-word recognition, ORF, phonological awareness, comprehension, writing, English and Maths at the beginning of Grade 1 and the end of Grades 1, 2, 4 and 7 (Taylor et al., 2017). The learners were tested in their HL which was Setswana (Taylor et al., 2017). The aim of this exercise was to assess the impact of the three different interventions. Once the survey instruments were completed, the data was captured and cleaned (Taylor et al., 2017).

Table 1: Summary of EGRS I data collection

Wave	**Grade**	**Year**
1	Beginning of Grade 1	2015
2	End of Grade 1	2015
3	End of Grade 2	2016
4	End of Grade 4	2018
5	End of Grade 7[1]	2020

The strength of EGRS I is that it is a panel dataset, and it tested the same basic reading components across multiple years. Regarding Randomised Control Trials (RCTs), their principal strength lies in their robust internal validity, which facilitates the precise measurement of causal impacts attributed to specific programs. In essence, if observed outcomes exhibit improvement in an intervention group, it can be confidently attributed to the intervention itself, allowing for a quantitative assessment of its influence. Nonetheless, it is important to acknowledge that RCTs are not without limitations. One notable drawback is their potential lack of external validity. While RCTs excel in ascertaining causal relationships within controlled settings, the applicability of findings to broader contexts can sometimes be limited. The controlled conditions under which RCTs operate might differ

[1] Of the reading components that I explored in this study, Grade 7s were only tested in ORF. I decided not to include this wave in my analysis for two reasons. One, I wanted to focus on the earlier grades and two, Wills et al. 2022a have already explored Grade 7 ORF in their paper.

significantly from real-world scenarios, raising questions about the generalisability of their results to diverse populations and conditions.

3.2 Description of variables

In my study, I used two sets of variables from EGRS I. The first set includes the reading component variables, and the second set includes the gender and grade variables. These are described below.

3.2.1 Reading component variables

Letter sounds

In the first three waves, learners were given an identical list of 110 simple sounds, and they were asked to sound as many as they could in a minute. The fieldworker recorded the total number that they sounded correctly. In the fourth wave, they were given another list of 110 sounds but this one included simple and complex consonant sequences (sounds such as ngw, thw). Wills et al. (2022a) explained that complex consonant sequences are more difficult to sound than simple sounds. There were four letter sound variables, each one representing the total number of sounds the learner sounded correctly in the wave.

Word recognition

In the first three waves, learners were given an identical list of 50 Grade 1 to 3 words (such as ‘fa’ which means give and ‘ja’ which means eat) and they were asked to read them without a time limit. The fieldworker recorded the total number that they read correctly. In the fourth wave, they were given a 70-word list to read. 21 of these words were included in the first list and the rest were a combination of Grade 1 to 3 words (such as ‘tlhapi’ which means fish and ‘tsamaya’ which means walk) and Grade 4 words (such as ‘tshireletso’ which means security or safety and ‘ditlhongwanamorago’ which means suffixes). There were four word recognition variables, each one representing the total number of words the learner read correctly in the wave.

Non-word recognition

In the second and third waves, learners were given an identical 50-word list of nonsense words. Non-words, which are made-up words that do not exist, cannot be memorised beforehand. Dubeck and Gove (2015, p. 317) explain that the non-word subtask "measures the ability to apply the knowledge of letter-sound relationships to decode unfamiliar words". There were two non-word recognition variables, each one representing the total number of non-words the learner read correctly in the two waves.

Oral Reading Fluency (ORF)

In the second, third and fourth waves, learners were given a short text and they were asked to read aloud as much of it as they could in a minute. The field worker recorded the number of words they read correctly in a minute. The difficulty, content and length of the text differed in each wave. In wave 2, learners were given a 62-word text about two children playing, washing and eating fruit. In wave 3, learners were given the same story with an additional sentence. There were 66 words in total. In wave 4, learners were given two texts. The first one had 159 words and it was about a dove and an ant and the second one had 216 words and it was about a hippo. There were four ORF variables, each representing the total number of words that the learner read correctly in a minute in each text.

Wave 4 comprehension scores

In wave 4, learners were given three minutes to read text 1, after which the fieldworker asked them a maximum of eight questions, dependent on how far the learner read up until in the text. Each of the questions was worth half a point, one point or two points and the maximum number of points a learner could earn was eight. There was one variable which represents the number of points that the learner achieved for comprehension.

Phonological awareness

Learners were tested on phonological awareness in the first three waves. Each test was different. In the first wave test, the first section included six questions and the second and third sections included three questions each. These are described in Table 2.

Table 2: Description of the wave 1 phonological awareness test

Section	Description	Total number of questions/words	Example
1.	Break up words into their sound components	6	roka
2.	Provide a different word starting with the same two sounds as the word provided	3	rata
3.	Provide a different word ending with the same two sounds as the word provided	3	leba

In wave 2, the learners were asked four of the questions that were administered in wave 1. These questions were from the first section and the words that they were asked to break up into their sound components were rre, podi, seme and mongwe. In wave 3, learners were asked three questions. These are described in Table 3

Table 3: Description of the wave 3 phonological awareness test

Section	Description	Total number of questions/words	Example
1.	Identify the position of a sound in the words provided (beginning, middle or end)	2	lesedi and roka
2.	Join two syllables together to form one word.	1	“phu” + “tha”

In the analysis stage, the scores in each wave were converted into percentages. The four common questions in the wave 1 test were also isolated so that I could compare them to the wave 2 results.

3.2.2 Other variables

Gender

Learners were asked their gender in every wave. Some of the results were a mismatch across waves. (I.e. in wave 1 the learner said that they were a boy but in wave 2 they said they were a girl). I used the binary variable which was the DBE's best estimate of gender.

Grade

Learners were asked what grade they were in every wave. In wave 1, every learner was in Grade 1. In wave 2, learners were in 1, 1A, 1B, 1C, 1D, 1E, 99 or "" and so because I was not sure what 99 and "" stood for and some of these learners were in Grade 2 in the next wave, I assumed that every learner was still in Grade 1. In wave 3, learners were either in Grade 1 or Grade 2. In wave 4, learners were in Grades 2, 3 or 4. I used the waves 3 and 4 grade variables in my analysis.

The descriptive statistics of the above-mentioned variables are presented in Table 4 below.

Table 4: Descriptive statistics of variables

Variable	Obs	Mean	Std. Dev	Min	Max
W1 total letter sounds	4538	5,1	9,8	0	99
W2 total letter sounds	4143	22,7	21,8	0	110
W3 total letter sounds	3781	39,5	26,4	0	110
W4 total letter sounds	3301	41,1	20,0	0	105
W1 total word recognition	4538	1,9	5,2	0	50
W2 total word recognition	4143	6,9	9,9	0	50
W3 total word recognition	3781	19,4	17,0	0	50

W4 total word recognition	3301	30,5	17,7	0	70
W2 total non-word recognition	4143	4,5	8,1	0	50
W3 total non-word recognition	3781	14,4	13,8	0	50
W2 ORF	4144	8,0	14,1	0	64
W3 ORF	3781	25,6	24,6	0	66
W4 ORF (text 1)	3301	48,9	33,5	0	159
W4 ORF (text 2)	3301	56,4	39,6	0	220
W4 comprehension score (text 1)	3301	2,5	1,8	0	8
W1 phonological awareness score (12 questions)	4538	18,1	25,9	0	100
W1 phonological awareness score (4 questions)	4538	20,3	33,0	0	100
W2 phonological awareness score (4 questions)	4143	17,5	29,4	0	100
W1 phonological awareness score (3 questions)	3781	60,8	35,8	0	100
Female (Dummy variable)	4514	0,5	0,5	0	1
W3 grade	3726	1,8	0,4	1	2
W4 grade	3301	3,7	0,5	2	4

3.3 Methods

In this section, I will describe the methods that I used to analyse the EGRS I data. In my analysis, I combined the control and treatment groups in the same way that Wills et al. (2022a) did. Their reason for not distinguishing across groups was that the impact sizes of the intervention were typically small. I used Stata and Excel to analyse the data. I will describe the methods I used to explore each research question below.

3.3.1 Performance in the fourth year of schooling

The first part of the study aimed to assess how learners performed in various reading components after four years of schooling. Of the five components that I wanted to explore, letter sound, word recognition and ORF were the only ones that were tested in wave 4.

Hence, in this section, I assessed these three components. I also assessed the performance in the wave 4 comprehension test. To assess how the 3301 learners who wrote the wave 4 test performed in these sections, I utilised pie charts, kernel densities (henceforth referred to as k-densities) and tables.

- Pie charts were used to assess the percentage of learners who were repeaters and those that were not, and they were also used to assess the percentage of all learners (including repeaters) as well as just the Grade 4s (excluding repeaters) who achieved each Grade 4 comprehension score.
- K-densities were used to explore the wave 4 performance in letter sounds, word recognition and ORF in texts 1 and 2. First I performed the analysis with repeaters and then I excluded them.
- Tables were used to record the percentage of learners who reached the DBE letter sound and ORF benchmarks. I analysed the percentages for all learners (including repeaters) as well as for each grade (Grade 4s or repeaters who were still in Grade 2 or 3). I also recorded the results separately for ORF texts 1 and 2 as opposed to using an average.

3.3.2 Pace of progress in reading components in the first four years of schooling

In the second section, I wanted to explore the pace of progress in reading components in the first four years of schooling. Of the components that I wanted to explore, all five were tested in at least two waves in EGRS I. To assess how learners performed, I used k-densities, line graphs and bar graphs.

- K-densities were used to explore the performance in letter sounds and word recognition in all four waves, in non-words in waves 2 and 3 and in ORF in waves 2, 3 and 4 (both texts). First, I performed this analysis including repeaters and then I excluded them. Each k-density figure included the results in every wave so that it is possible to compare the performance across waves.

- Line graphs were used to record the percentage of learners reaching the DBE letter sound and ORF benchmarks. First, I performed the analysis including repeaters and then I excluded them. This allows us to compare performance across years of schooling and grades of schooling. I decided to average the percentage who reached the ORF benchmarks in text 1 and 2 so that the trajectory of the learning profile was clearer. Another line graph was used to display the percentage of all learners and Grade 4s who recognised more than 10 words in each wave.
- Bar graphs were used to record the percentage of learners achieving each phonological awareness score for the four-common-question-test that was administered in waves 1 and 2.

3.3.3 Gender performance in reading components

In the third section, I wanted to explore whether there was a difference in gender performance in the five reading components. I used pie charts, k-densities, line graphs and bar graphs.

- Pie charts were used to record the proportion of each gender who passed (more than 50 percent), failed or did not get any of the Grade 4 comprehension questions correct.
- K-densities were used to compare the performance in letter sounds, word recognition, non-word recognition and ORF in each wave. Repeaters were included in the analysis and then they were excluded.
- Line graphs were used to track the percentage within each gender who reached the DBE benchmarks in each wave. Repeaters were included and then they were excluded in the analysis and the performance in texts 1 and 2 was averaged.
- Bar graphs were used to compare the gender performance in the wave 1 and 2 four-question and wave 3 three-question phonological awareness tests.

3.3.4 Relationship between Grade 4 comprehension scores and reading components

In section 4 I wanted to explore the relationship between reading components in each wave and Grade 4 comprehension scores. I performed kernel-weighted local polynomial regressions of every reading component on Grade 4 comprehension scores. Avery (2013, p. 1) defines local polynomial regression as "a nonparametric technique for smoothing scatter plots and modelling functions". I used the lpoly command (with the default options) in Stata which has the option of including 95 percent confidence bands. Repeaters were included in the analysis.

3.4 Limitations of EGRS I

Unfortunately, there are several limitations with the EGRS I data. The first is that it is not nationally representative. This means that we cannot make generalisations about the performance in South Africa as a whole or even for non-fee or North West learners. The second is that the phonological awareness tests that were administered in the first three waves were different. The third is that the letter sound and word recognition tests that were administered in wave 4 were more difficult versions of the tests administered in the first three waves. The fourth is that the ORF text was different in every wave. The wave 2 and 3 tests were 95 percent similar, however. The fifth limitation is that non-words were not tested in wave 1 and wave 4, ORF was not tested in wave 1 and phonological awareness was not tested in wave 4. The sixth limitation is that fieldworkers may have made mistakes in the data collection and the DBE may have made mistakes while cleaning the data. The seventh is that the recorded gender of some learners changed across waves. The eighth is that the grade variables in the first two waves were not clear. My assumption that the wave 2 learners who wrote the test were in Grade 1 might be incorrect. The ninth limitation is attrition. Quite a number of learners who participated in the study in wave 1 left the school or were absent on the day of testing in the next three waves. For example, 4538 learners wrote the letter sound test in wave 1, 4143 wrote it in wave 2, 3781 wrote it in wave 3 and 3301 wrote it in wave 4. This means that 27 percent of the original sample was not included in wave 4. The tenth limitation is that learners were not tested in Grade 3. This means that

we do not know how much learning took place in Grade 3 and how much took place in Grade 4.

CHAPTER FOUR: DATA PRESENTATION AND DISCUSSION

4.1 Introduction

In 2017, the tip of the comprehension iceberg was revealed when the PIRLS 2016 results were published. It was discovered that 78 percent of Grade 4s could not read for meaning in any language. In recent years, however, many researchers have dived under the surface of the water and explored the bottom of the iceberg - the basic reading components that researchers have argued lead to comprehension. Today, we know a lot more about the bottom of the iceberg. The most notable contributions are the discovery that isiZulu and siSwati learners in Mpumalanga and isiXhosa learners in the Eastern Cape are entering schools with underdeveloped phonological awareness and listening skills (Wills et al., 2022b); Setswana and isiXhosa learners are only slowly mastering letter and complex consonant sequences (Wills et al., 2022b); learners across the Nguni and Sotho-Setswana languages have low and slow ORF trajectories and girls perform better than boys (Wills et al., 2022b); there is a positive relationship between letter sounds and ORF and ORF and comprehension for Nguni and Sotho-Setswana languages (Pretorius & Spaull, 2022; Wills et al., 2022b). Moreover, the DBE has developed letter sound and ORF benchmarks for the Nguni, Sotho-Setswana, EFAL and Afrikaans languages (Ardington et al 2020, 2021 and 2022; Mohohlwane et al., 2022, Wills et al., 2022b) and calculated the percentage reaching these targets when they are supposed to (Mohohlwane et al., 2022).

In this chapter of the paper, I aimed to contribute to the existing base by exploring more of the iceberg below the surface. I used EGRS I data to assess how learners performed in various reading components in the fourth year of schooling as well as how they progressed in these components across four years of schooling. It would be useful to know, for example, if the majority of learners reach the Grade 1 letter sound benchmark in Grade 2, as opposed to Grade 1 as it suggests that most of the learning is taking place in the second year. In this chapter, I also assessed whether there is a gender gap at the basic reading component level (and not just at the comprehension level) and I explored the relationship between performance comprehension scores in the fourth year of schooling and the pace of progression in reading components in the earlier years.

The results chapter is made up of four sections. In the first section, I explored how the EGRS I learners performed in different reading components after four years of schooling. In the second section, I dove deeper by analysing how these learners performed in various reading components in the years leading up to Grade 4. I used learning profiles to capture the change in performance across waves. In the third section, I disaggregated these learning profiles by gender to assess whether girls outperformed boys in the way that the literature suggests. In the fourth and final section, I explored the relationship between Grade 4 comprehension scores and reading components in the earlier grades.

4.2 Performance in the fourth year of schooling

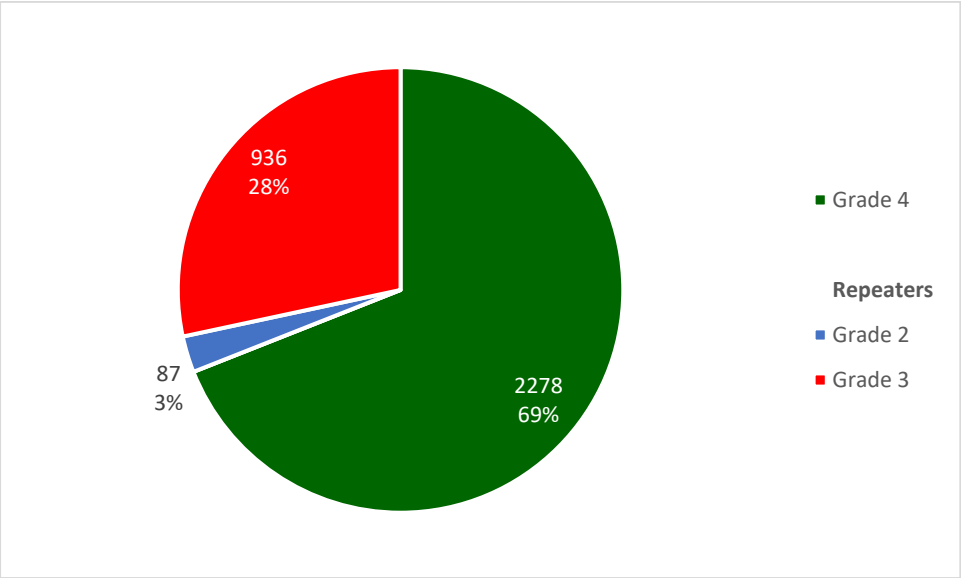
In this first section, I explored how EGRS I learners who were taught in Setswana performed in various reading components after at least four years of schooling. Referring back to the iceberg theory mentioned in the literature section, we can see the tip of the iceberg and know from PIRLS 2016 that 78 percent of Grade 4 learners cannot read for meaning in any language. PIRLS 2016 also revealed that 90 percent of learners participating in the Setswana test could not read for meaning (Howie et al., 2017). This section aims to increase our knowledge of the bottom of the iceberg. The bottom of the iceberg is made up of reading components such as phonological awareness, letter sounds, ORF and vocabulary. If learners are struggling with any of these components, there is a high chance they will not be able to read with comprehension. Hence, it is useful to assess how the EGRS I learners performed in these components after four years of schooling.

In wave 4, 3301 learners were tested in letter sounds, word recognition and ORF and so these are the components I focused on in this section. Non-words and phonological awareness were not tested in wave 4, hence these components were only dealt with in the next three sections.

First, I assessed how all learners who wrote these tests in wave 4 performed and then I disaggregated the results by grade. After four years of schooling, the learners should have been in Grade 4, however, this was not the case because some learners were held back one or two years. In wave 1, all learners were in Grade 1 but in wave 4 as can be seen in Figure

1, 87 learners were in Grade 2, 936 were in Grade 3 and 2278 were in Grade 4. It is misleading to assess all learners together because some of those learners were exposed to the Grade 3 and 4 curriculums whereas others were not. This may explain why some learners did not progress as much. I therefore analysed the performance after four years of schooling and I also analysed the performance after two, three and four grades of schooling.

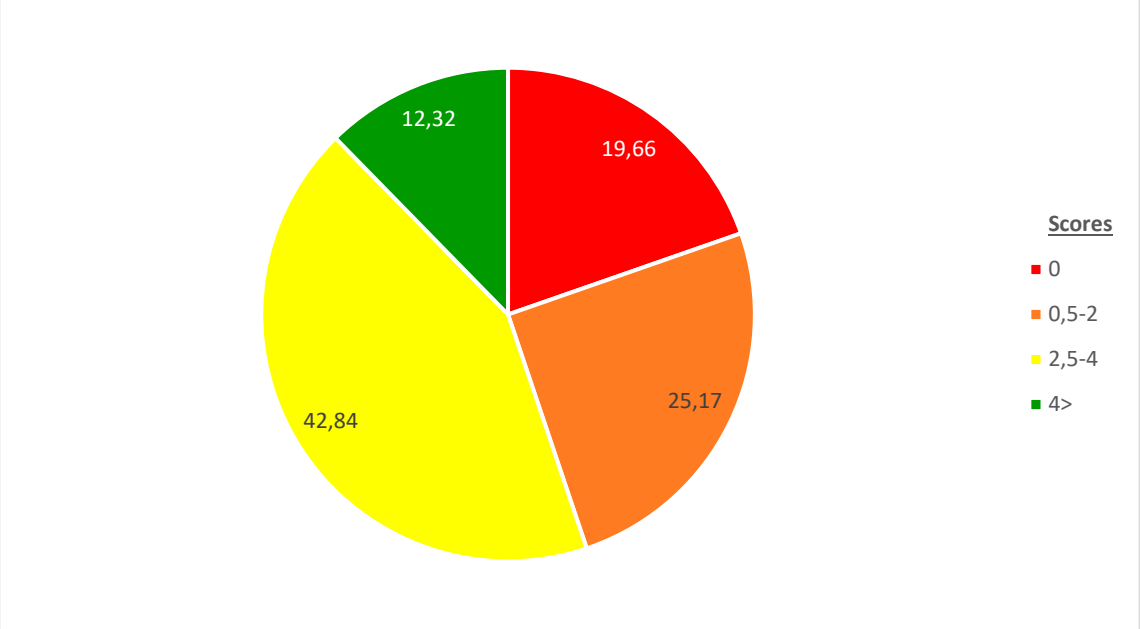
Figure 1: Percentage of learners in each grade in wave 4



As part of the wave 4 test, learners were given three minutes to read a 159-word story about a dove and an ant. At the end of the three minutes, field workers tested whether the learners understood the story by asking them a maximum of eight questions. The number of questions that were asked was dependent on how far the learner read up until in the text. Each of these questions was worth half a point, one point or two points and the maximum number of points a learner could earn was eight.

Figure 2 is a pie chart displaying the percentage of learners achieving a range of scores. 19,7 percent of the 3301 learners who wrote the Grade 4 test achieved a score of zero, 68 percent achieved between half a point and 4 and 12,3 percent achieved a score greater than 4. Stated differently, 87,7 percent of learners achieved a score of four or less and 1 in 5 learners could not answer a single question correctly. These results confirm the PIRLS 2016 finding that 90 percent of learners who wrote the Setswana test could not read for meaning.

Figure 2: Percentage of learners achieving a range of comprehension scores (including repeaters)



As is expected, the percentage of learners who achieved a high score increased when repeaters were excluded from the sample. This can be seen in Figure 3. Of the 2278 learners in Grade 4 who wrote the test, 9,5 percent achieved a score of zero, 73,2 percent achieved a score between half a point and 4 and 17,3 percent achieved a score greater than 4. While it is slightly reassuring that the percentage of learners achieving a score of zero decreased from 19,7 to 9,5 when repeaters were excluded, it is significant that one in ten Grade 4 learners could not answer a single question correctly and eight in ten Grade 4 learners failed the test. In the rest of this section, I will explore how these learners performed in the basic reading components that were tested in wave 4.

Figure 3: Percentage of learners achieving a range of comprehension scores (excluding repeaters)

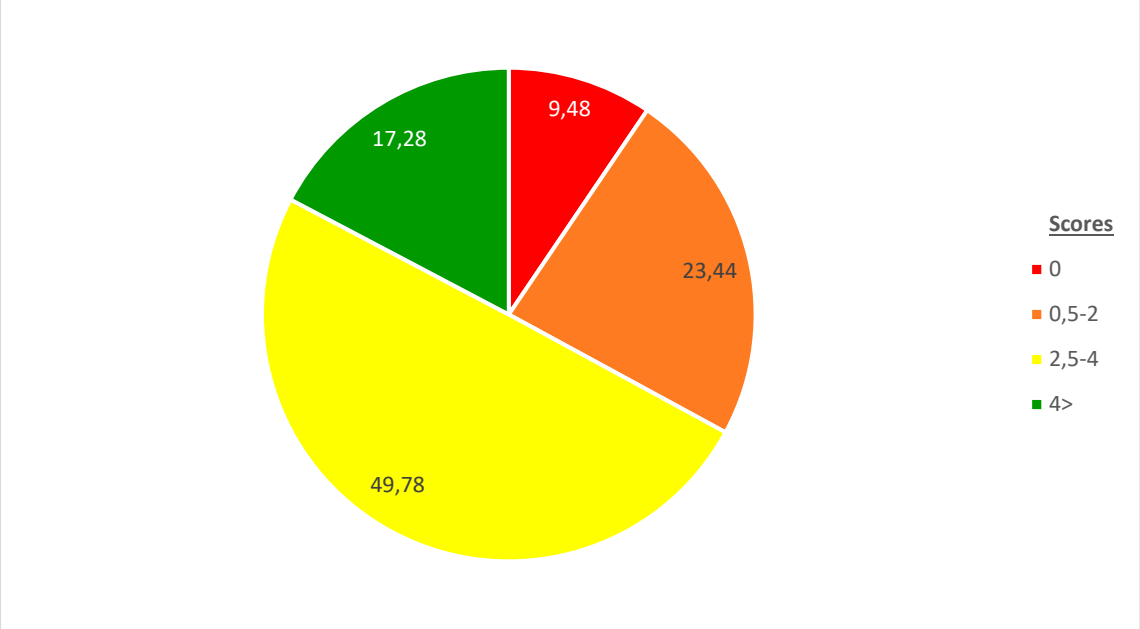


Figure 4 is a k-density of the distribution of the number of letters learners who wrote the wave 4 test read correctly. This figure included repeaters. As part of the wave 4 test, learners were given a list of 110 sounds and they were asked to sound as many as they could in a minute. Unlike the previous three letter sound tests, this one included complex consonant sequences (such as ngw, thw) which as explained by Wills et al. (2022a) are more difficult to sound. The end of Grade 1 letter sound benchmark that the DBE has set is 40 LCPM and this is represented by the red vertical line. Learners who reach this benchmark stand a better chance of reading with comprehension by the end of Grade 4.

This figure along with Table 5 revealed that just under half of the learners (46,5 percent) had not reached the Grade 1 letter sound benchmark after at least four years of schooling and were therefore over three years behind. As can be seen in Figure 5, this image improves slightly when the learners are divided into grades. In this case, 61,4 percent of Grade 4s (those who did not repeat a grade in the four years), 37,8 percent of Grade 3s (those who repeated one grade in the four years) and 16,1 percent of Grade 2s (those who repeated two grades in the four years) reached the Grade 1 letter sound benchmark. This indicates that the learners with a weaker letter sound ability were held back, however, it is still significant that 38,6 percent of those who did not repeat a grade in the four years did not reach the benchmark, and 4 percent could not read more than 10 letters. By Grade 4,

learners should be reading in English but if learners cannot master one of the first early grade reading components in their HL, it is unsurprising that the reading comprehension levels are so low. These results need to be viewed with caution, however, as the wave 4 letter sound test included complex consonant sequences that are more difficult. If it only included simple sounds, the percentage reaching the benchmark likely would have been higher.

Figure 4: K-density of the number of letter sounds learners read correct in a minute in wave 4 (including repeaters)

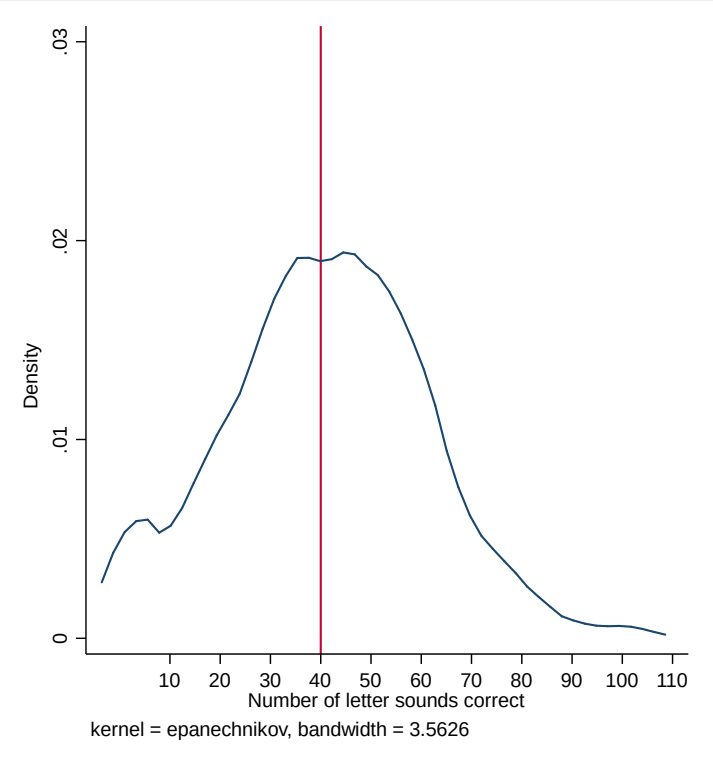


Figure 5: K-density of the number of letter sounds learners in each grade read correctly in a minute after four years of schooling

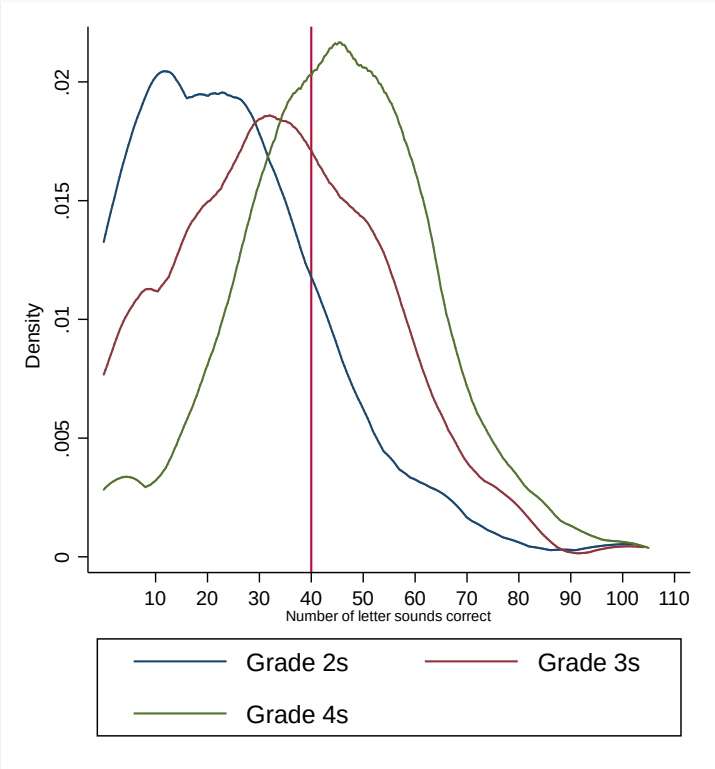


Table 5: Percentage of learners reaching the Grade 1 letter sound benchmark after four years of schooling.

	Grade 1 benchmark
All learners	53,5
Grade 2 (repeaters)	16,1
Grade 3 (repeaters)	37,8
Grade 4	61,4

In wave 4, learners were also given a list of 70 words that they were asked to read. This list included simple words such as 'ja' which means 'eat' and more difficult words such as 'setlhongwanapele' which means 'prefix'. Figure 6 is a k-density of the distribution of the number of words learners recognised. This distribution is bimodal with a large peak at approximately 40 words and a smaller one at 5 words. When repeaters are included, 15,7 percent of learners could not read more than 5 words, 20,2 could not read more than 10

words and 28,7 could not read more than 20 words. As can be seen in Figure 7, the repeaters are responsible for the smaller peak. 67,8 percent of the Grade 2s and 38,3 percent of Grade 3s recognised less than 10 words compared to 9,3 percent of Grade 4s. While 10 percent is still a fair portion of learners, there was a group of learners in Grade 4 who made progress in word recognition. 47,4 percent of the Grade 4s recognised between 35 and 50 words out of a total of 70 words.

Figure 6: K-density of the number of words recognised in wave 4 (including repeaters)

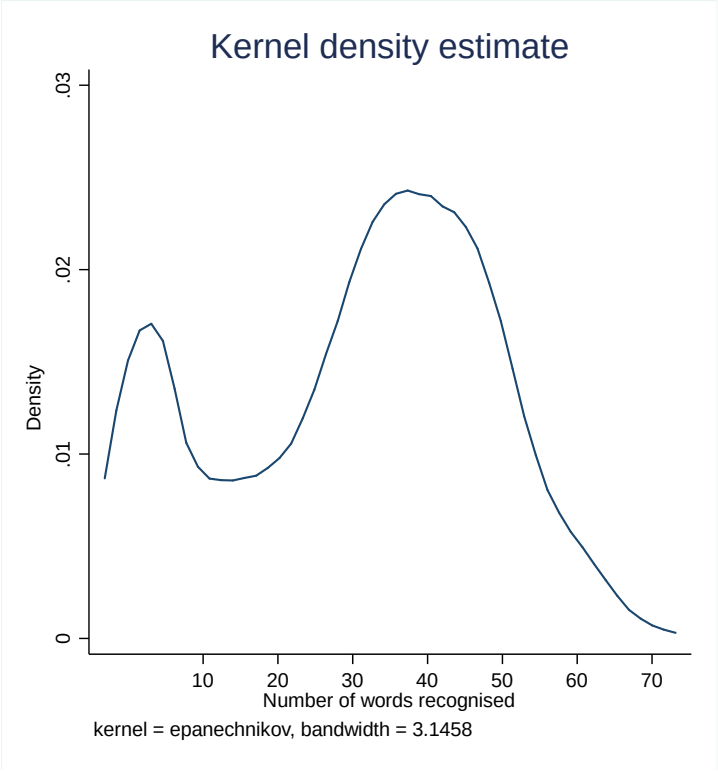
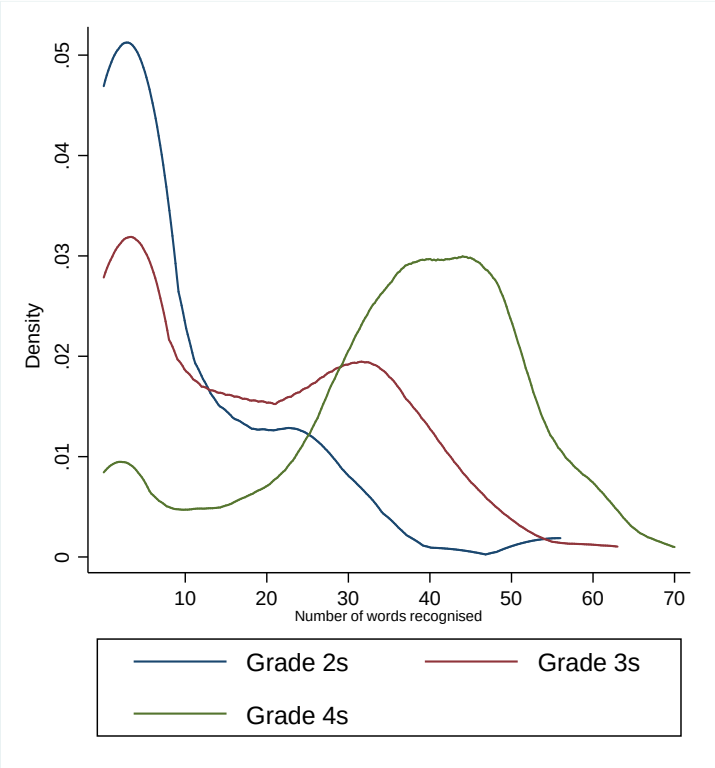


Figure 7: K-density of number of words recognised in each grade after four years of schooling.



As mentioned already, in the wave 4 test, learners were given a 159-word text about a dove, but they were also given a second 216-word text about a hippo. They were given one minute to read each text and the fieldworker recorded how many words they read correctly. The recorded figure is known as ORF. Figures 8 and 10 are k-densities of the distribution of ORF for texts 1 and 2 respectively. The end of Grade 2 and 3 benchmarks that the DBE has set, and which are represented by the two vertical red lines are 40 and 60 WCPM respectively. Learners who reach these benchmarks when they should stand a better chance of reading with comprehension by the end of Grade 4. Even though the texts were different, the shape of each k-density is similar. Like word recognition, these k-densities are bimodal with a larger peak at approximately 60 WCPM and a smaller one at 10 WCPM. In text 1, 14,1 percent of learners could not read any words and 18,4 percent read less than 10 words. As can be seen in Table 6, 62 percent reached the Grade 2 benchmark and 39 percent reached the Grade 3 benchmark. These figures were similar for text 2 with the exception that 46,8 percent of learners reached the Grade 3 benchmark. Expectedly, and as can be seen in Figures 9 and 11, the smaller peak is driven by the repeaters. 31,2 and 27,3 percent of Grade 2 and 3 learners combined could not read a single word in texts 1 and 2 respectively. In comparison, 6,5 and 5,7 percent of Grade 4s could not read a single word in each text. There

was a group of learners in Grades 3 and 4, however, who made greater progress in ORF. Between 29,8 and 36,8 percent of Grade 3s and between 77,4 and 81,3 percent of Grade 4s reached the Grade 2 benchmark and between 51,5 and 61,2 percent of Grade 4s reached the Grade 3 benchmark. It is notable, however, that between 40 and 50 percent of Grade 4s did not reach the Grade 3 benchmark a year after they were supposed to, and 20 percent did not reach the Grade 2 benchmark two years after they were supposed to.

Figure 8: K-density of ORF for text 1 in wave 4 (including repeaters)

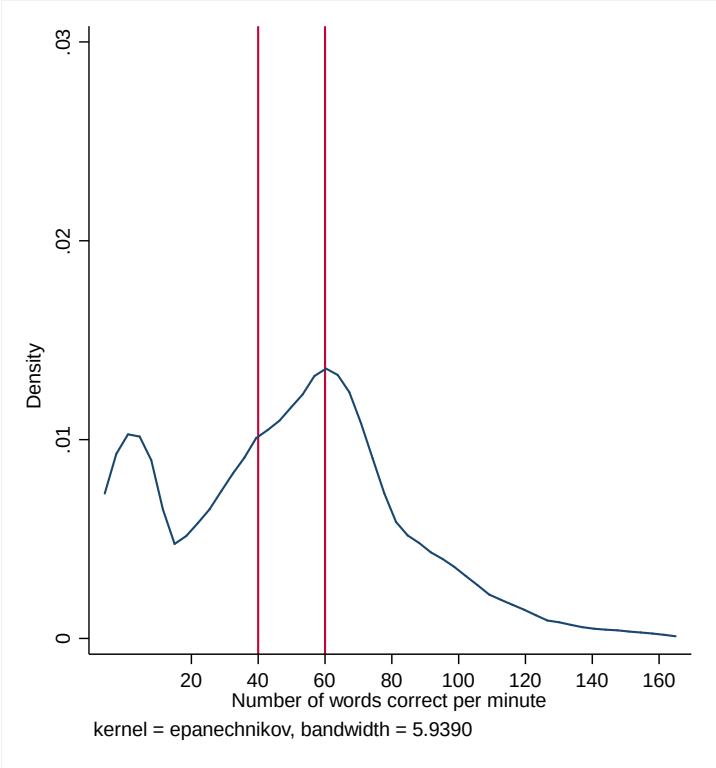


Figure 9: K-density of ORF for text 1 in each grade after four years of schooling.

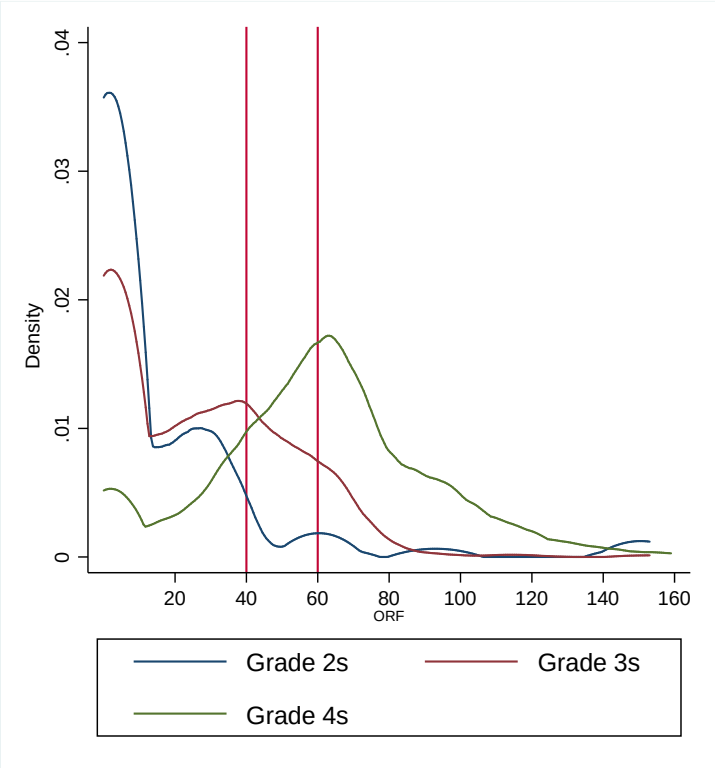


Figure 10: K-density of ORF for text 2 in wave 4 (including repeaters)

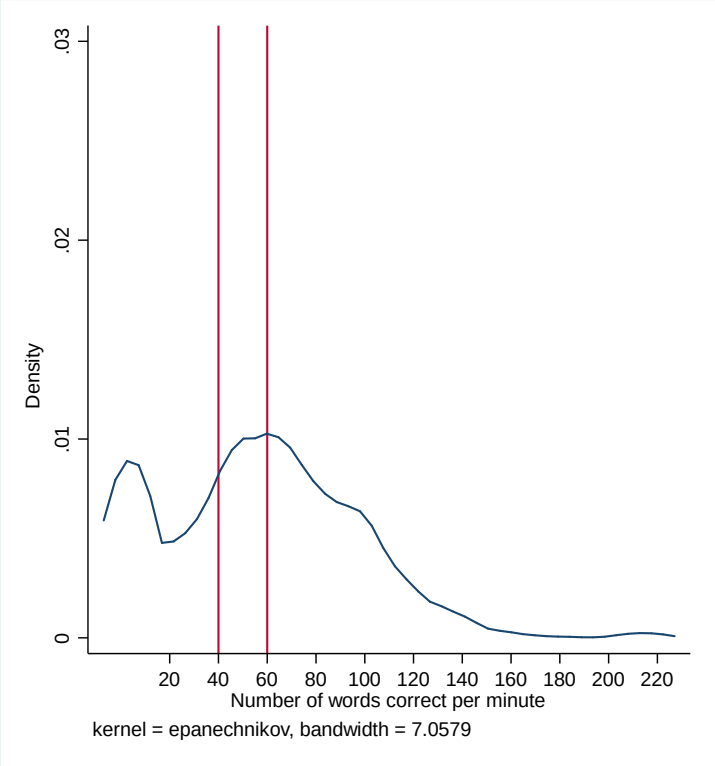


Figure 11: K-density of ORF for text 2 in each grade after four years of schooling.

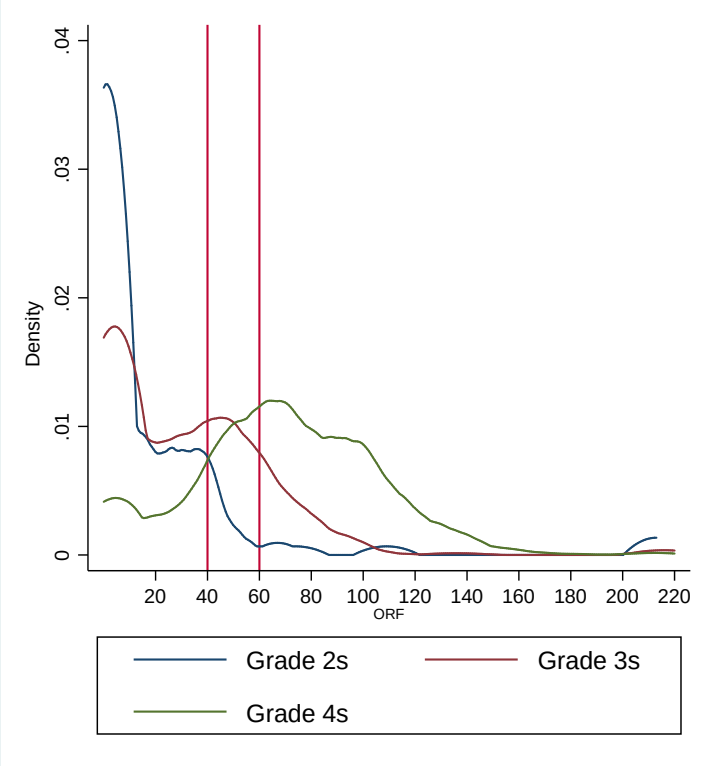


Table 6: Percentage of learners reaching the Grade 2 and 3 benchmarks after four years of schooling.

	Grade 2 benchmark (40 WCPM)	**Grade 3 benchmark (60 WCPM)**
All learners	62 - 66,8	38,7 - 46,8
Grade 2 (repeaters)	6,9 - 10,3	4,6 - 5,8
Grade 3 (repeaters)	29,8 - 36,8	10,8 - 15,5
Grade 4	77,4 - 81,3	51,5 - 61,2

In this first section, I have confirmed the importance of exploring the iceberg underneath the surface of the ocean. It is not enough to know that 78 percent of learners cannot read for meaning in any language and 90 percent cannot read for meaning in Setswana. We need to assess how they are performing in basic reading components. The first finding is that after four years of schooling, 54 percent of learners could not reach the Grade 1 letter sound benchmark. This number decreases when repeaters are excluded but it is still significant that 39 percent of Grade 4s could not sound more than 40 letters correct in a

minute. In other words, four in ten children after four years of school had not mastered the most basic building block of reading, letter sound recognition. The second finding is that 20,2 percent of learners could not recognise more than 10 out of a total of 70 words after four years of schooling and 28,7 percent could not recognise more than 20 words. Again, the repeaters are the ones responsible for these high numbers, however, it is significant that 10 percent of Grade 4s could not recognise more than 10 words. The third finding is that there is a group of learners who made progress in word recognition. 37 percent of learners could recognise between 35 and 50 words after four years of schooling and this number increases to 47 percent when the repeaters are excluded. The fourth finding is that there are four groups of ORF learners: non-readers (those who could not read a single word), very slow readers (those who could read at least one word but did not reach the Grade 2 benchmark), slow readers (those who reached the Grade 2 benchmark but not the Grade 3 benchmark) and on-track learners (those who reached the Grade 3 benchmark). As measured by text 1, 14 percent of learners were non-readers, 24 percent were very slow readers, 23 percent were slow readers and 39 percent were on-track. When repeaters are excluded, 6 percent of Grade 4s were non-readers, 16 percent were very slow readers, 26 percent were slow readers and 51 percent were on-track. While half of the Grade 4s reached the Grade 3 benchmark, 45 percent made progress but not enough for comprehension and 5 percent did not make any progress at all.

4.3 Pace of progress in reading components in the first four years of schooling

In this second section, I built on the findings in the first section by exploring the pace of progress in basic reading components in the first four years of schooling. To do this, I compared the performance in reading component tasks across the four waves. In the first section, I only explored letter sounds, word recognition and ORF, but in this section, I extended this list by including non-words and phonological awareness, as these two components were tested in the earlier waves. The order in which I explored the components was letter sounds, word recognition, non-word recognition, ORF and phonological awareness. It is important to note that each of the waves has a different sample size due to attrition.

In the first three waves, learners were given an identical list of 110 simple sounds to read, and the fieldworker counted how many letters they sounded correct in a minute. As mentioned in the previous section, they were given a list of 110 simple sounds and complex consonant sequences in wave 4. When comparing the four waves, it is important to keep in mind that the test in the fourth wave was more difficult as it included complex consonant sequences. Figures 12 and 13 display the k-densities of the distribution of the number of letters learners sounded correct in the first four waves.

At the beginning of Grade 1, 41,3 percent of the learners could not sound a single letter correctly. Since a large percentage of learners entered Grade 1 without any letter sound skills, it suggests that learners were either not exposed to letter sounds in the Early Child Development (ECD) phase, they were exposed to it, but it was not taught properly, or they were not enrolled in an ECD centre. The percentage of learners who could not sound a single letter dropped to 13,3 percent after one year of schooling, it dropped further to 8,5 percent after two years of schooling and it dropped even further to 2,2 percent after four years of schooling. This indicates that the majority of learners are exposed to this skill in school. The wave 2 and 3 k-densities are both bimodal. After one year of schooling, there was a large peak at approximately 5 LCPM and a smaller one at the Grade 1 letter sound benchmark (40 LCPM) and after two years of schooling there were two peaks with similar heights at approximately 5 LCPM and 50 LCPM. Since the 5 LCPM peak was smaller after two years of schooling, it means that many of the learners who could not read more than 5 LCPM at the end of Grade 1 made greater progress over the next year. As discussed already, the k-density after four years of schooling is approximately normally distributed with a large peak at 40 LCPM and a tiny one at 5 LCPM. This peak is smaller than the peak after two years of schooling which means that even more alphabetically illiterate learners made progress over the next two years. However, since the larger peak was pushed back from 50 LCPM to 40 LCPM between the second and fourth year of schooling, it suggests that some learners performed worse between the two years. This could be due to a text effect as the wave 4 test was more difficult or it could be due to changes in the sample (attrition). 4538 learners wrote the letter sound test in wave 1 compared to 3301 in Grade 4 which is a 27 percent attrition rate. If the learners who left the school or were absent on the day of testing achieved better scores, this would explain the leftward shift.

In Figure 14, I display the percentage of learners who reached the Grade 1 letter sound benchmark in each year of schooling. At the beginning of school, 1,7 percent of learners reached the benchmark, after one year of schooling 23,9 percent of learners reached the benchmark, after two years of schooling 52,3 percent of learners reached the benchmark and after four years of schooling, 53,5 percent of learners reached the benchmark. This indicates that less than a quarter of the learners reached the benchmark when they were supposed to, and half of the learners reached it a year later. Moreover, between the second and fourth year of schooling, the learning profile is flat, suggesting that after two years of schooling, learners who had not mastered their letter sounds did not get the opportunity to do so. This plateau could be explained by the more difficult nature of the wave 4 test though as learners likely struggled with the complex consonant sequences. When repeaters are excluded in waves 3 and 4, as can be seen in Figure 15, 58,6 and 61,4 percent of Grade 2 and 4 learners reached the benchmark respectively which is an improvement despite the profile remaining flat.

Figure 12: K-densities of letters correct per minute in the first four waves (including repeaters)

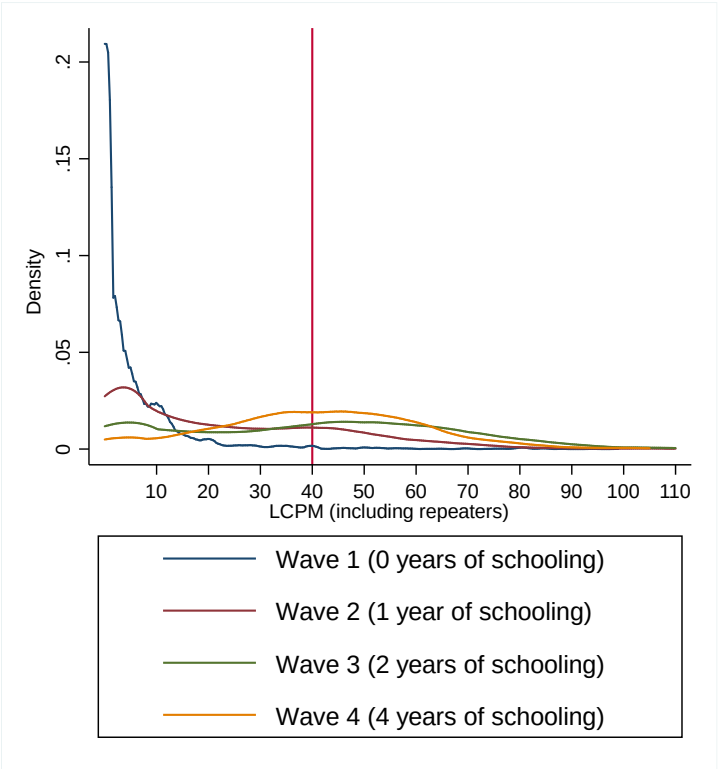


Figure 13: K-densities of letters correct per minute in waves 2, 3 and 4 (including repeaters)

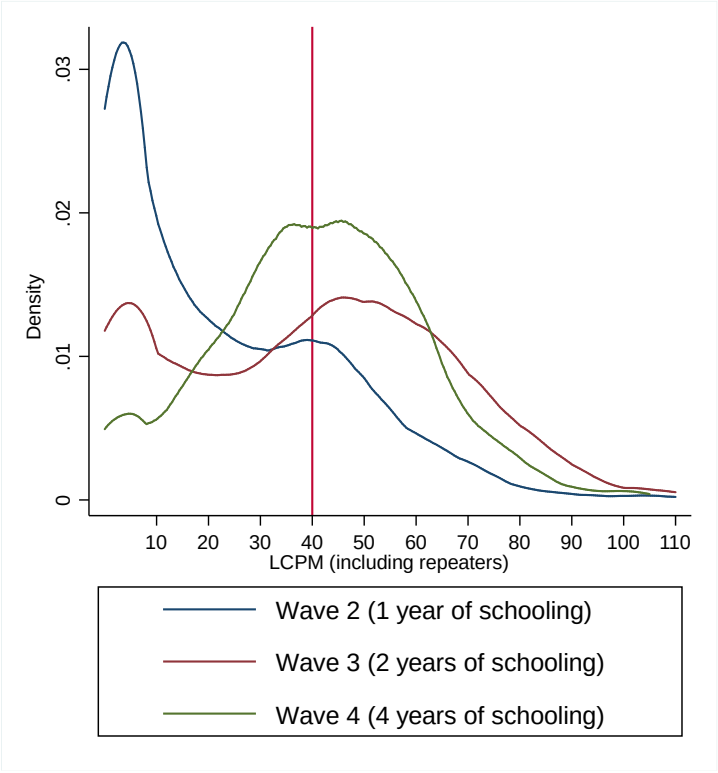


Figure 14: Learning profile displaying the percentage of learners reaching the DBE Grade 1 letter sound benchmark in each wave (including repeaters)

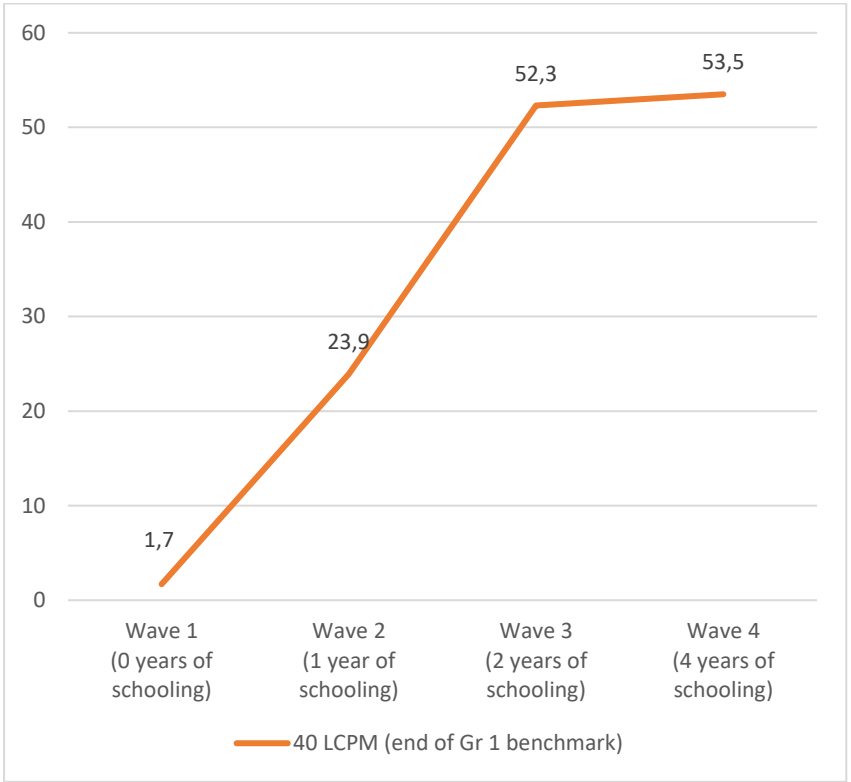
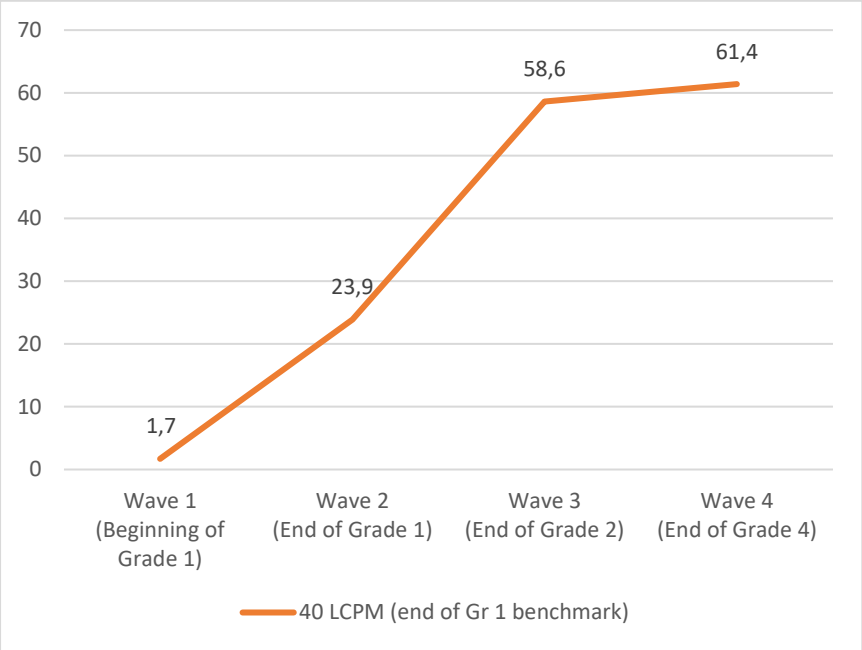


Figure 15: Learning profile displaying the percentage of learners reaching the DBE Grade 1 letter sound benchmark in each wave (excluding repeaters)



In the first three waves, learners were given an identical 50-word list of words and the fieldworker counted the number that they read correctly. As mentioned in the previous section, in wave 4, learners received a 70-word list. 21 of these words were from the previous 50-word list, and the rest were a combination of simple and more difficult words. Figures 16 and 17 display the k-densities of the distribution of the number of words learners recognised in the first four waves. At the beginning of school, 65,4 percent of the learners could not recognise any words. This is unsurprising as learners are only expected to learn how to read words in Grade 1. After one year of schooling, this number had dropped significantly to 29,7 percent but as can be seen in Figure 19 only 21,9 percent of learners could read more than 10 words. After two years of schooling, the percentage of learners who could not recognise a single word dropped to 15,7 and 57 percent of learners were evenly distributed across 11 and 50 words. In Figure 20, the graph is flatter between waves 1 and 2 than between waves 2 and 3, an indication that a greater percentage of learners are learning to read in the second year of schooling. At the end of Grade 1, approximately 20 percent of learners could read more than 10 words, but this figure tripled over the next year. As discussed already, the percentage of learners who could not recognise any words after four years of schooling dropped to 6 percent, and the percentage who could read

more than 10 words increased to 79,8 percent. There was also a greater percentage of learners reading between 30 and 50 words.

As expected, when repeaters are excluded as can be seen in Figure 18, the percentage of learners who could not recognise more than 10 words decreased from 43 (with repeaters) to 35,6 percent at the end of Grade 2 and from 20,2 to 10 percent in Grade 4. This means that in Figure 19, the gradient between waves 2 and 3 is even steeper, indicating that an even greater proportion of learners learnt how to read more than 10 words in Grade 2. All in all, though, the pace of progress in word recognition was slow (especially in Grade 1) and many learners passed through grades without this important skill.

Figure 16: K-densities of word recognition in the first four waves (including repeaters)

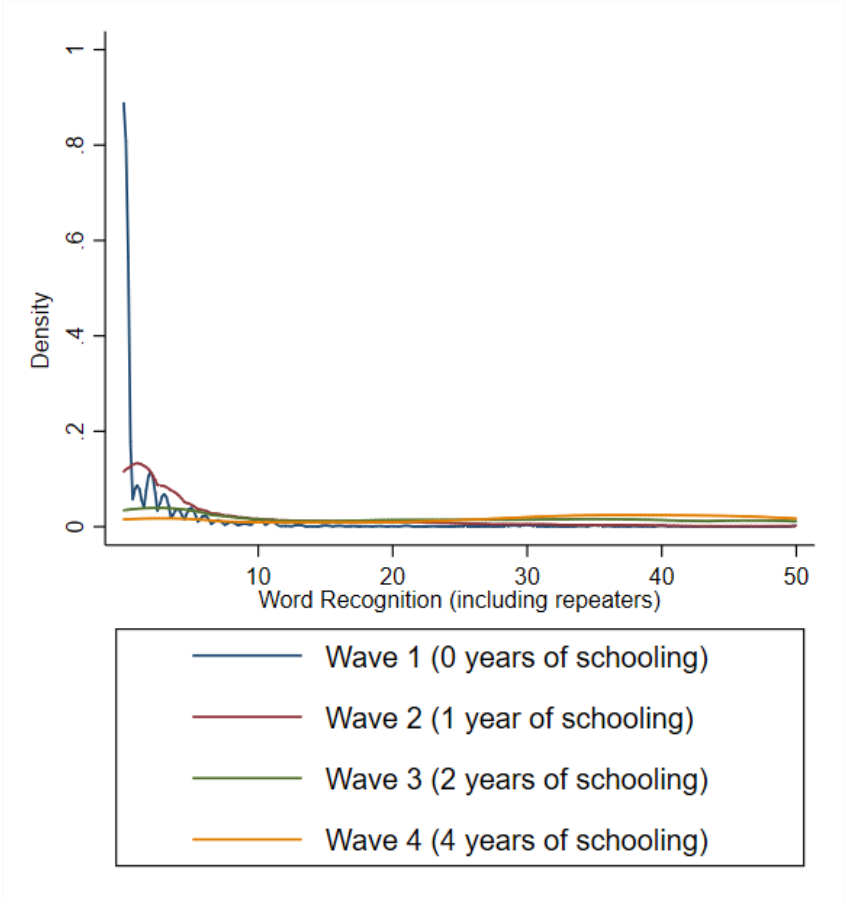


Figure 17: K-densities of word recognition in waves 2, 3 and 4 (including repeaters)

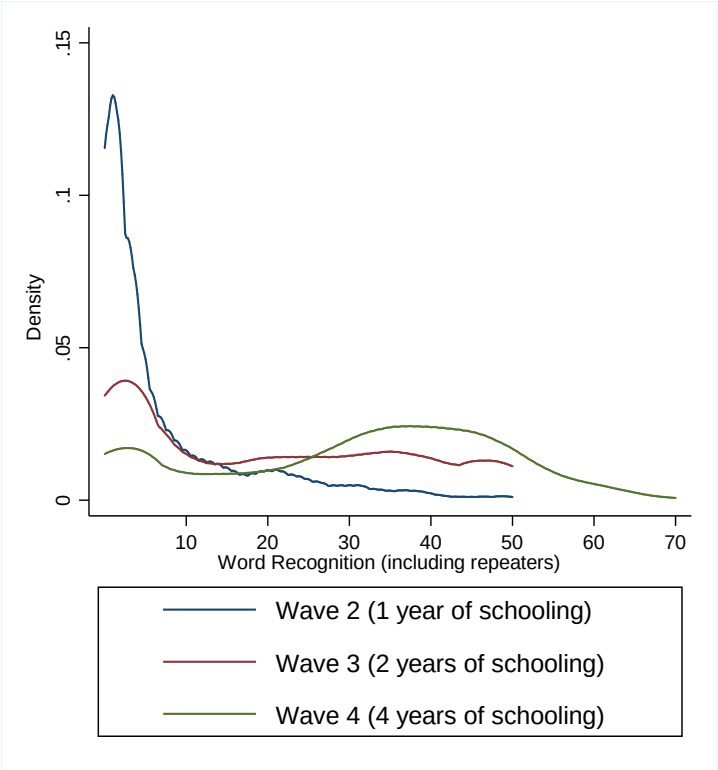


Figure 18: K-densities of word recognition in waves 2, 3 and 4 (excluding repeaters)

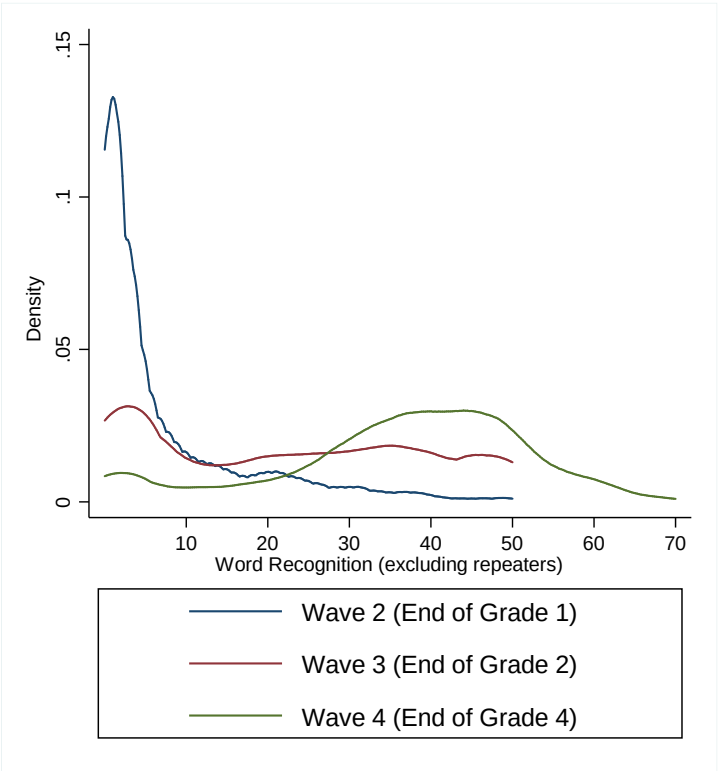
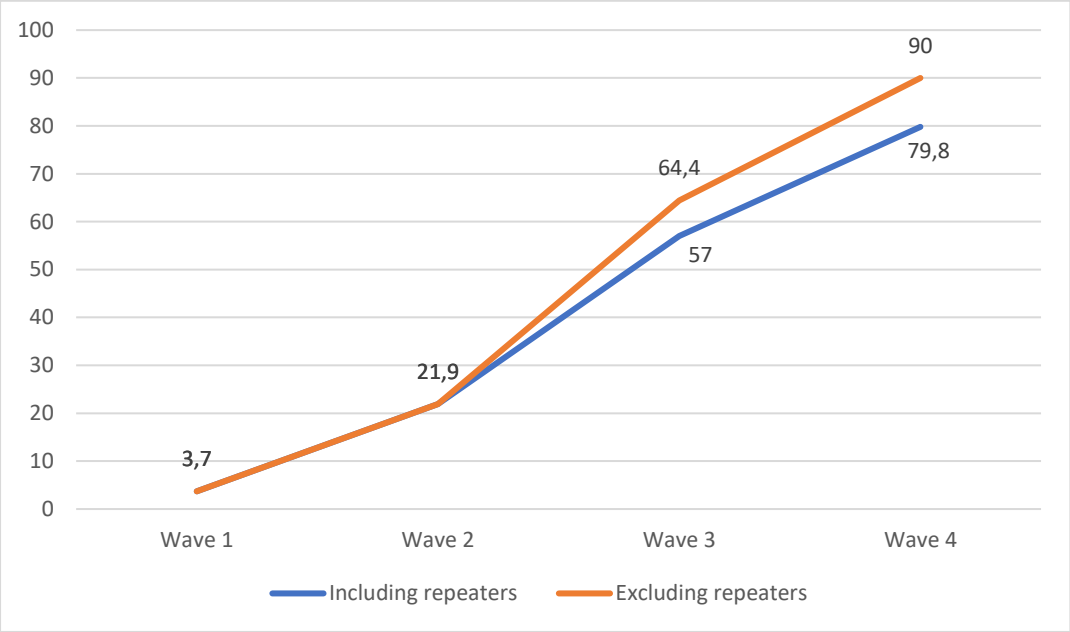


Figure 19: Learning profiles displaying the percentage of learners who could recognise more than 10 words in each wave (including and excluding repeaters)



Non-words are made-up words that can be sounded out using phonics. Similar to the word recognition subsection, learners were given an identical 50-word list in waves 2 and 3 and asked to read them without a time limit. Figure 20 displays the k-densities for the distribution of the number of non-words learners recognised in waves 2 and 3. After one year of schooling, the vast majority of learners (65,6 percent) could not recognise any non-words. After two years of schooling, this number dropped to 33,9 percent and a greater percentage of learners recognised approximately 30 non-words. When repeaters are excluded in wave 3, as can be seen in Figure 21, the percentage who could not recognise any non-words dropped further. These results reveal that by the end of Grade 1, the majority of learners could not read a single non-word, and over the next year, only slow progress was made. After two years of schooling, 45,6 percent of learners were still unable to read more than 10 non-words and when repeaters are excluded, this figure drops to 38,6. The poor performance in non-words is unsurprising, given the poor knowledge of letter sounds.

Figure 20: K-densities of non-word recognition in waves 2 and 3 (including repeaters)

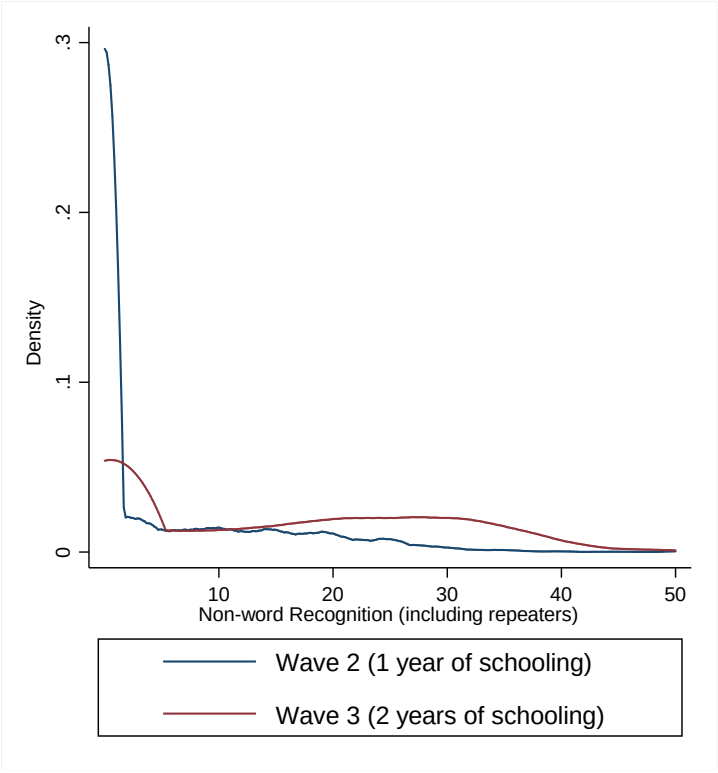
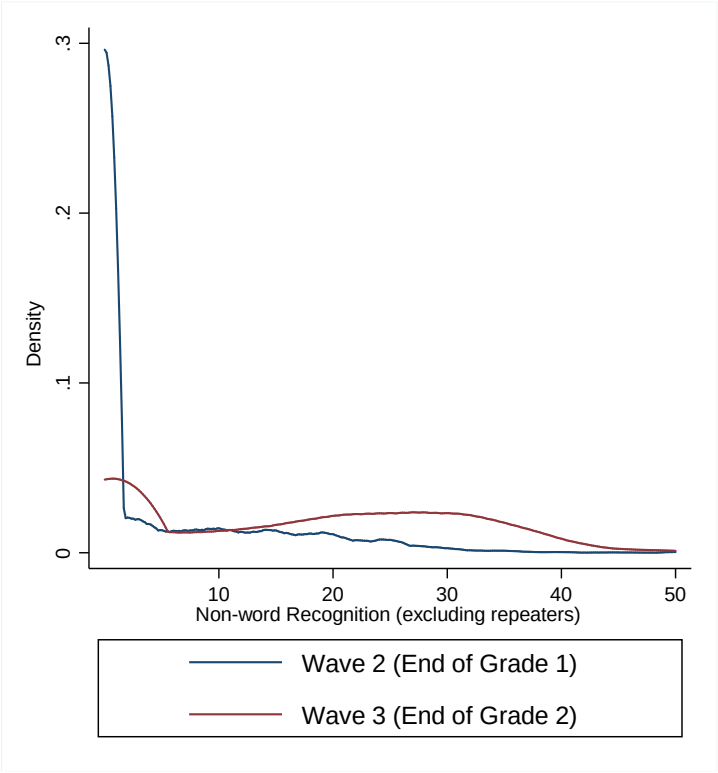


Figure 21: K-densities of non-word recognition in waves 2 and 3 (excluding repeaters)



Figures 22 and 23 display the k-densities for the distribution of ORF in waves 2, 3 and 4. The red vertical lines represent the DBE Grade 2 and Grade 3 benchmarks which are 40 WCPM and 60 WCPM respectively. In each wave (bar wave 1), learners were given a minute to read a simple passage of text and in wave 4, they were given two passages. The fieldworker recorded the number of words the learner read correctly. The difficulty, content and length of the text differed in each wave and therefore the results across years cannot be directly compared with one another, but nonetheless, it is still useful to assess the general trend. After one year of schooling, a large percentage (60,5 percent) of learners could not read a single word correctly in a minute. There was, however, a smaller peak at 20 WCPM. After two years of schooling, the percentage of learners who could not read a single word decreased to 36,2 percent and there were five smaller peaks (each one larger than the previous one) from approximately 20 WCPM to 60 WCPM. This suggests that after two years of schooling, a small group of learners started to develop the skill. As has already been discussed, in wave 4, learners were given two texts. The k-density of the first text is bimodal with a small peak at approximately 10 WCPM and a larger one at the Grade 3 benchmark (60 WCPM). The k-density of the second text is similar although there was a greater percentage of learners who read more than 80 WCPM. 14,1 and 12,4 percent could not read a single word in text 1 and text 2 respectively which is better than the performance in wave 3.

Figure 24 displays the percentage of learners who reached the Grade 2 and 3 benchmarks in each wave. 5,5 percent of learners reached the Grade 2 benchmark after one year of schooling, 35,4 percent reached it after two years of schooling and between 62 and 66,8 percent reached it after four years of schooling. In addition, 0,9 percent of learners reached the Grade 3 benchmark after one year of schooling, 12,3 percent reached it in after two years of schooling and between 38,7 and 46,8 percent reached it after four years of schooling. This means that two in three learners did not reach the Grade 2 benchmark when they were supposed to and one in three learners did not reach it two years later. Moreover, over half of the learners did not reach the Grade 3 benchmark a year after they should have. Both learning profiles have a positive gradient which means that learning is taking place, however since the slopes are flatter than they should be, the pace that learning is occurring is much too slow.

When repeaters are excluded, as can be seen in Figure 25, the percentage of learners reaching the benchmarks in each grade increases. This is expected. 5,5 percent of learners reached the Grade 2 benchmark at the end of Grade 1, 41,5 percent reached it at the end of Grade 2 and between 77,4 and 81,3 percent reached it at the end of Grade 4. Furthermore, 0,9 percent of learners reached the Grade 3 benchmark at the end of Grade 1, 14,5 percent reached it at the end of Grade 2 and between 51,5 and 61,2 percent reached it at the end of Grade 4. Even though this is an improvement, it is still significant that three in five learners had not reached the Grade 2 benchmark by the end of Grade 2 and one in five learners had not reached it by the end of Grade 4. Moreover, at least two in five Grade 4s did not reach the Grade 3 benchmark a year after they should have.

Figure 22: K-densities of ORF performance in waves 2, 3 and 4 (including repeaters)

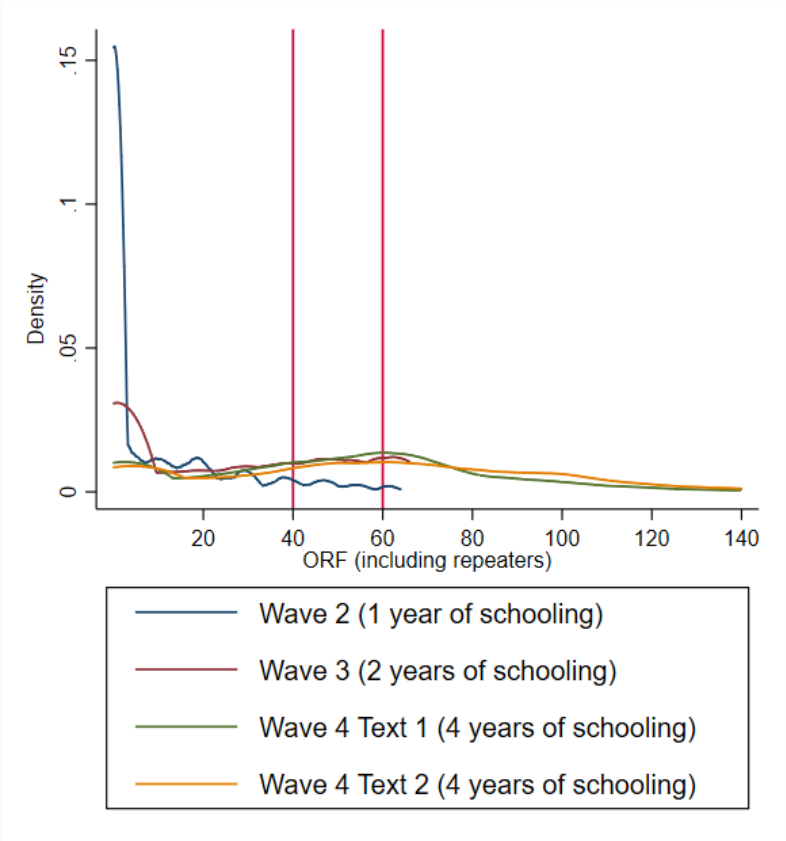


Figure 23: K-densities of ORF performance in waves 3 and 4 (including repeaters)

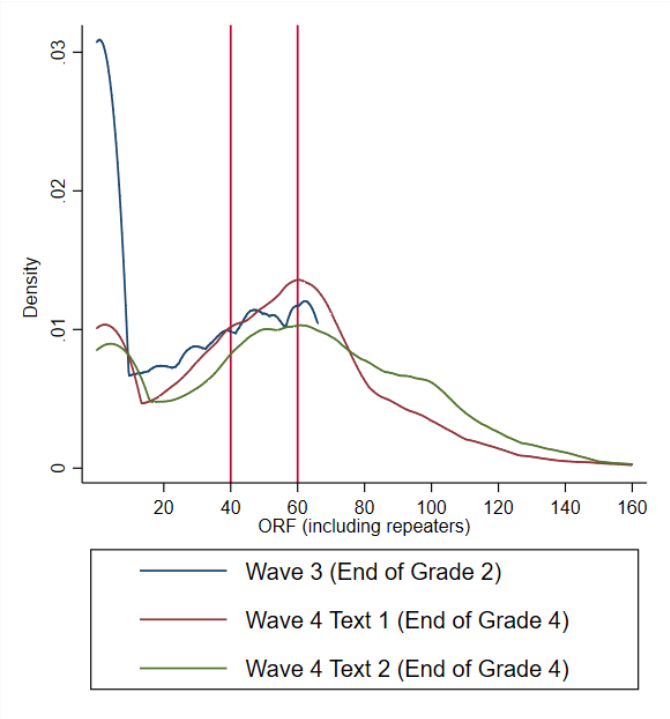


Figure 24: Learning profiles displaying the percentage of learners reaching the DBE Grade 2 and 3 benchmarks in each wave (including repeaters)

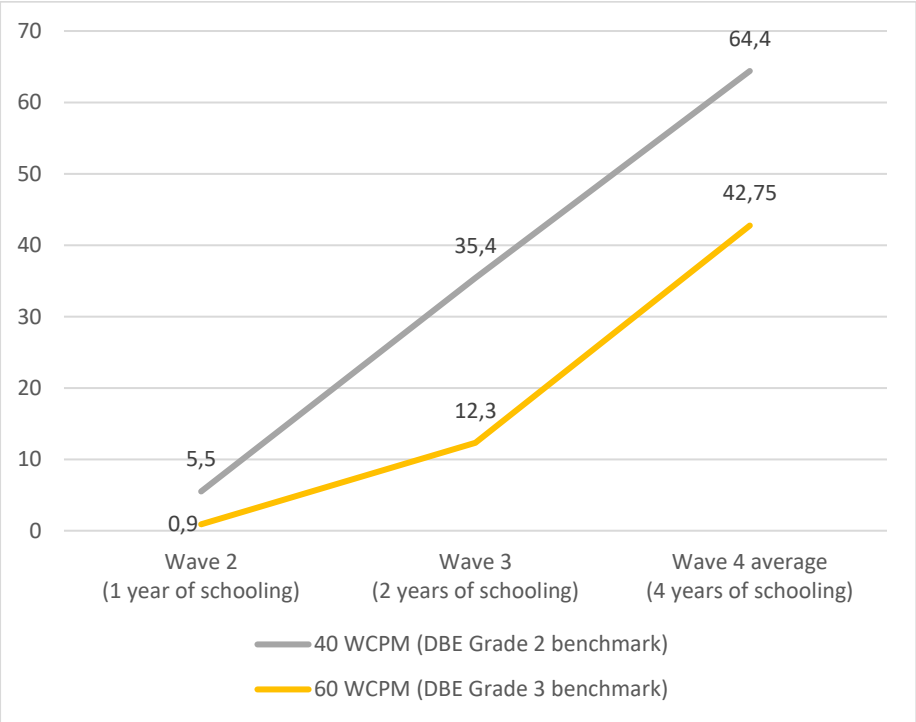
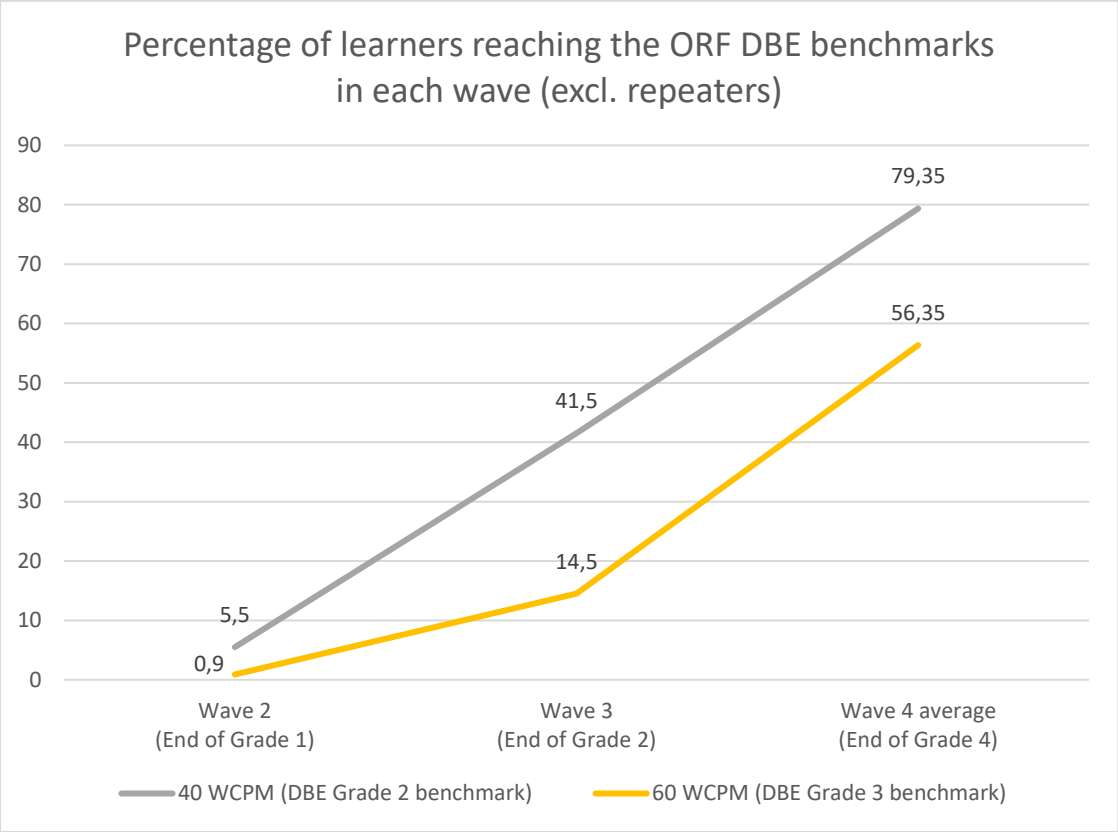


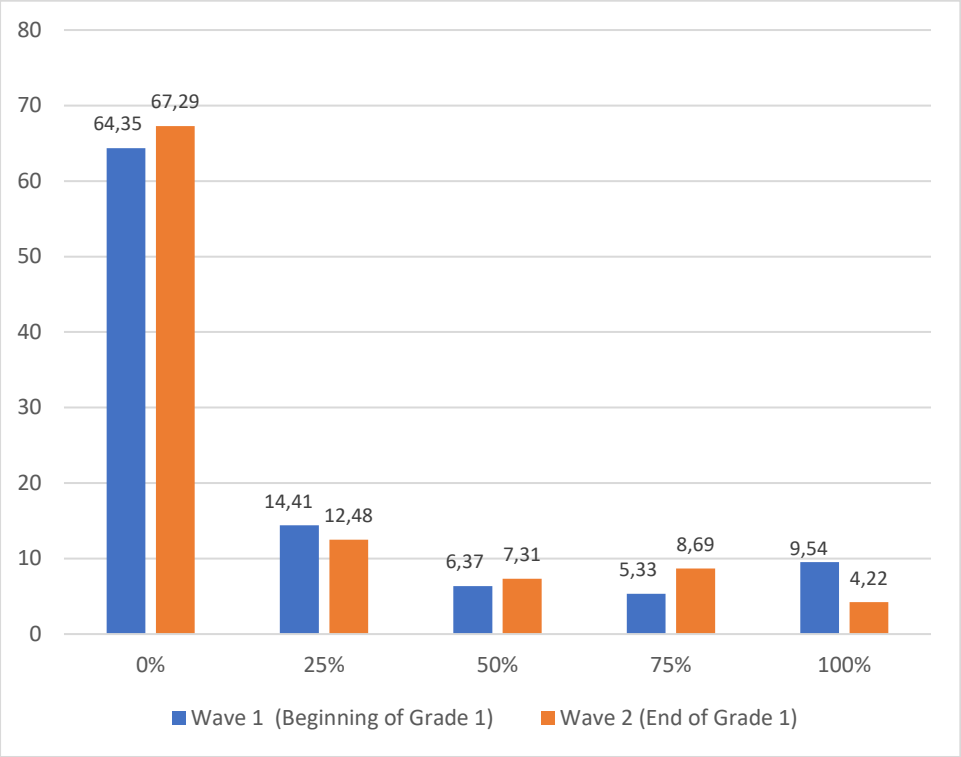
Figure 25: Learning profiles displaying the percentage of learners reaching the DBE Grade 2 and 3 benchmarks in each wave (excluding repeaters)



The phonological awareness tests that were administered in waves 1 to 3 were different so we cannot directly compare the scores, however, since four of the questions that were asked in wave 1 were asked in wave 2, we can directly compare this subsection. Learners in both waves were asked to break up four words into their sound components. Figure 26 reveals that at the beginning of Grade 1 64,4 percent of learners did not get a single question correct, 14,4 percent got one question correct, 6,4 percent got two questions correct, 5,3 percent got three questions correct and 9,5 percent of learners got all four questions correct. At the end of Grade 1, 67,3 percent of learners did not get a single question correct, 12,5 percent got one question correct, 7,3 percent got two questions correct, 8,7 percent got three questions correct and 4,2 percent of learners got all four questions correct. This means that by the end of Grade 1, there was a smaller percentage who achieved a perfect score and a larger percentage who could not get a single question correct. There are two reasons why these findings are troubling. The first is that nearly two in three learners entered Grade 1 with no phonological awareness skills. The second is that

over the next year, there was no improvement and some of the learners who had the skill (or at least a portion of it), lost it. This suggests that phonological awareness is not being taught properly in the ECD phase and Grade 1. The attrition rate between waves 1 and 2 was 8,7 percent, however. If learners with better scores left the school, this may explain the decrease in phonological awareness skills.

Figure 26: Percentage of learners achieving different phonological awareness scores at the beginning and end of Grade 1



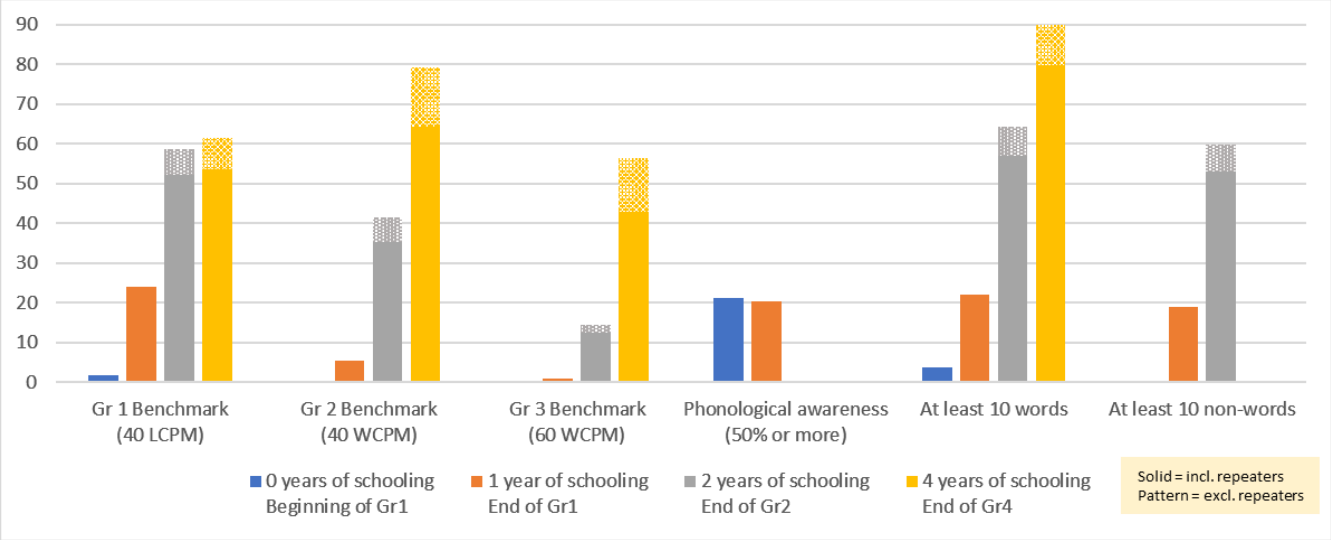
In this section, I explored how learners performed in the different reading components across the first four years of schooling. Figure 27 provides a snapshot of the performance in each of these reading components. I found that the majority of learners entered Grade 1 with poor phonological awareness and letter sound skills. By the end of Grade 1, a greater percentage of learners had made some progress in letter sounds, however, majority had not reached the Grade 1 letter sound benchmark. By the end of Grade 1, only 20 percent of learners reached it and by the end of Grade 2, just over half of the learners had reached it. Over the next two years, letter sound progress flatlined, suggesting that if learners did not know their letter sounds by the end of Grade 2, they did not receive the opportunity to learn them in the following two years. This flatline could be explained by the inclusion of

complex consonant sequences in the wave 4 test, however. It is also noteworthy that over a space of a year, learners did not improve their phonological awareness score and some learners lost a portion of skills that they did have. This suggests that learners were not taught phonological awareness skills in Grade 1.

By the end of Grade 1, learners had also made slow word recognition and non-word recognition progress. Four in five learners could not recognise more than 10 words and 65,6 percent could not recognise any non-words. The poor phonological awareness and letter sound performance could explain this troubling result. In the second year of schooling, learners seemed to make greater progress with word recognition. 15,7 percent of learners could not recognise a single word and only two in five learners could not recognise more than 10 words. After four years of schooling, two in ten learners could not recognise more than 10 words and when repeaters are excluded, this decreased to one in ten learners. Findings for non-word recognition are worse. After two years of schooling, nearly one-third of learners could not recognise any non-words and nearly half of the learners could not recognise more than 10 non-words.

The ORF progress is also slow. After a year of schooling, 60,5 percent of learners were non-readers and after two years of schooling, 36,2 percent of learners were non-readers. When repeaters are included two-thirds of learners did not reach the Grade 2 benchmark after two years of schooling and one-third of learners did not reach it two years later. This means that after two years of schooling, only one in three learners was on-track. As was discovered in the first section, only 40 percent of learners reached the Grade 3 benchmark after four years of schooling. All in all, the majority of learners made slow progress in these basic skills in the first four years of schooling, and some did not make any progress. This could explain why many learners are not prepared in Grade 4.

Figure 27: Percentage of learners achieving each reading component

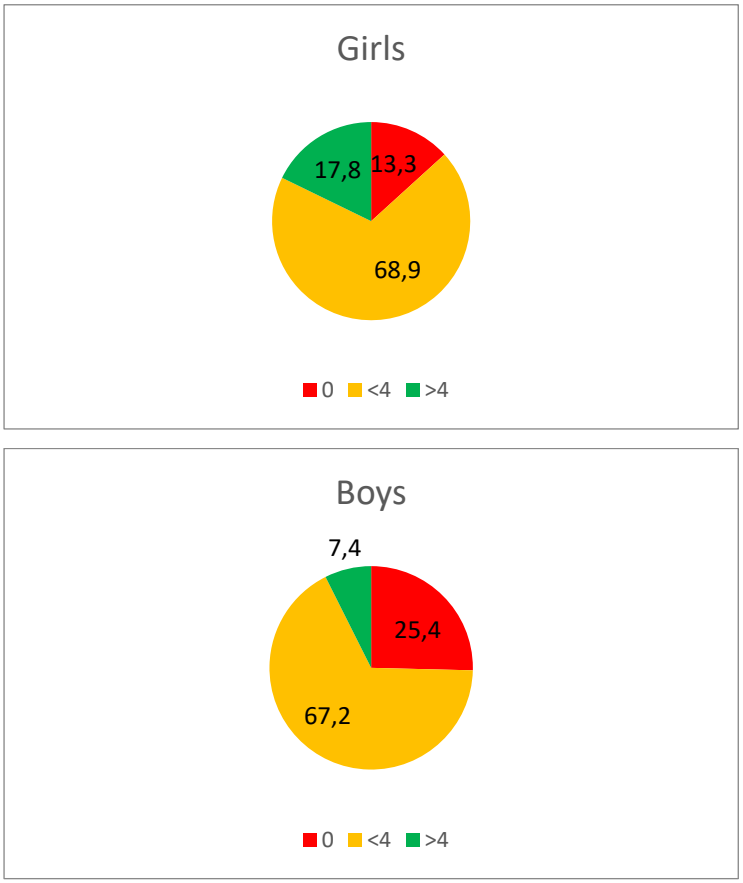


4.4 Gender performance in reading components

Several international, African and South African studies have found that in many countries, girls outperform boys in reading comprehension. According to the PIRLS 2016 results, in South Africa, 72 percent of girls could not read for meaning in any language compared to 84 percent of boys (Howie et al., 2017). In this section, I explored whether this trend held in EGRS I and I also analysed whether there were gender differences at the basic reading component level.

As explained in the first section, in wave 4, learners were given three minutes to read a 159-word story about a dove and an ant. Once the three minutes were up, field workers asked each learner a maximum of eight questions. The number of questions that were asked was dependent on how far the learner read up until in the text. Each of the questions was worth half a point, one point or two points and the maximum number of points a learner could earn was eight. As seen in Figure 28, 18 percent of girls passed the Grade 4 comprehension test compared to only 7 percent of boys and 13 percent of girls did not get any of the questions correct compared to 25 percent of boys. It is also notable that 25 percent of girls who wrote the wave 4 test were in Grade 2 or 3 compared to 36 percent of boys. This is an indication that boys may be falling off the bus within the first few years of schooling.

Figure 28: Percentage of boys and girls who passed, failed (between 0,5 and 4) or did not get any questions correct in the Grade 4 comprehension test (including repeaters)



In this third section, I will disaggregate the learning profiles by gender to assess whether this trend held at a basic reading component level. For each wave, I will compare the reading component k-densities for each gender. I will start with letter sounds and then I will compare the performance for word recognition, non-word recognition, ORF and phonological awareness.

Figure 29 compares the letter sound performance for each gender. The wave 1 k-densities (Figure 29a) have similar shapes, although a greater percentage of boys could not sound any letters, and a greater percentage of girls could sound approximately 8 and 20 LCPM. This indicates that from the beginning of Grade 1, there were already signs that girls were starting to outperform boys. Figure 29b compares the performance in wave 2. After one year of schooling, there was a larger percentage of boys who could read less than 10 LCPM and a greater percentage of girls who could read between 30 and 90 LCPM. This suggests that by the end of Grade 1, a greater percentage of girls reached the Grade 1 letter sound

benchmark. Figure 30 confirms this. 29,8 percentage of girls in wave 2 reached the benchmark compared to 18,5 percent of boys. In other words, 1,6 times as many girls reached the benchmark at the end of Grade 1.

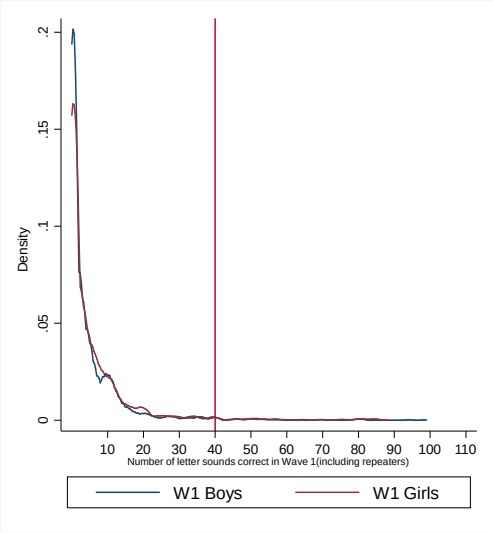
Figure 29c compares the performance in wave 3. After two years of schooling, there was a greater percentage of boys who did not reach the Grade 1 letter sound benchmark and there was a greater percentage of girls who exceeded the benchmark. Approximately the same percentage of girls and boys could read exactly 40 LCPM. Figure 30 confirms that 60,2 percent of girls reached the benchmark compared to 45,4 percent of boys. In this wave, there were 1,3 times as many girls as boys who reached the benchmark which means that the gap between genders in terms of the percentage reaching the benchmark was smaller than it was in wave 2.

Figure 29d compares the wave 4 performance. After four years of schooling, a larger percentage of boys could read between 0 and 30 LCPM and a greater percentage of girls could read between 30 and 100 LCPM. Figure 30 reveals that in this wave, 61,1 percent of girls reached the benchmark compared to 46,8 of boys.

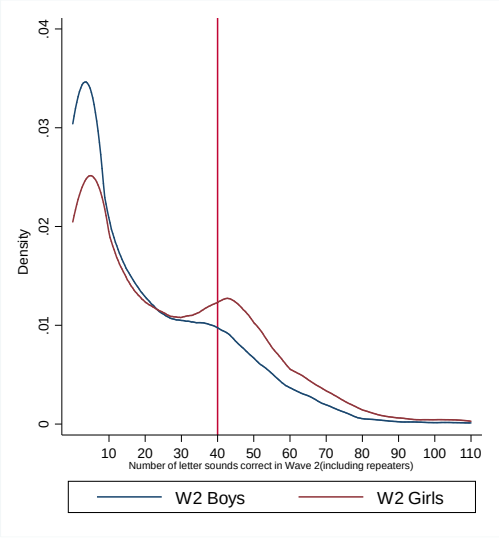
For the first three waves, whether repeaters are excluded or not, the learning profile for girls is steeper than the learning profile for boys. This means that for each year of schooling, girls learned more letter sounds than boys. The learning profiles for both genders between waves 3 and 4 are flat when repeaters are included but when repeaters are excluded, as can be seen in Figure 31 the learning profile for boys is slightly steeper. Since 62 percent of the repeaters in both waves were boys, it makes sense that the learning profiles for boys would be slightly steeper when repeaters are excluded than when they are included.

Figure 29: K-densities of letter sound performance in the four waves for each gender (including repeaters)

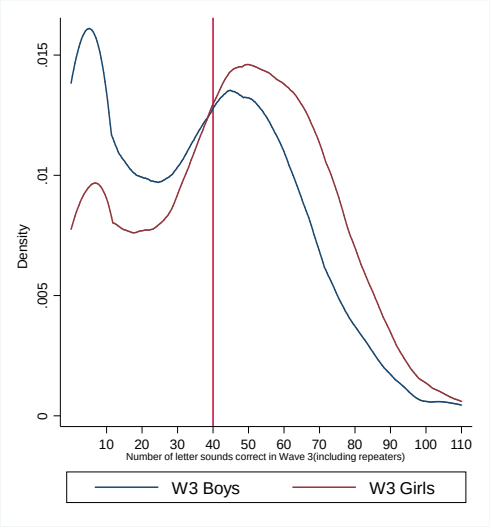
a (0 years of schooling)



b (1 year of schooling)



c (2 years of schooling)



d (4 years of schooling)

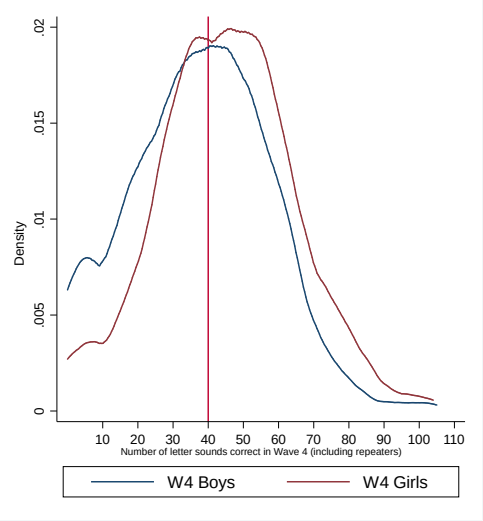


Figure 30: Percentage of learners in each gender group who reached the DBE Grade 1 letter sound benchmark in each wave (including repeaters).

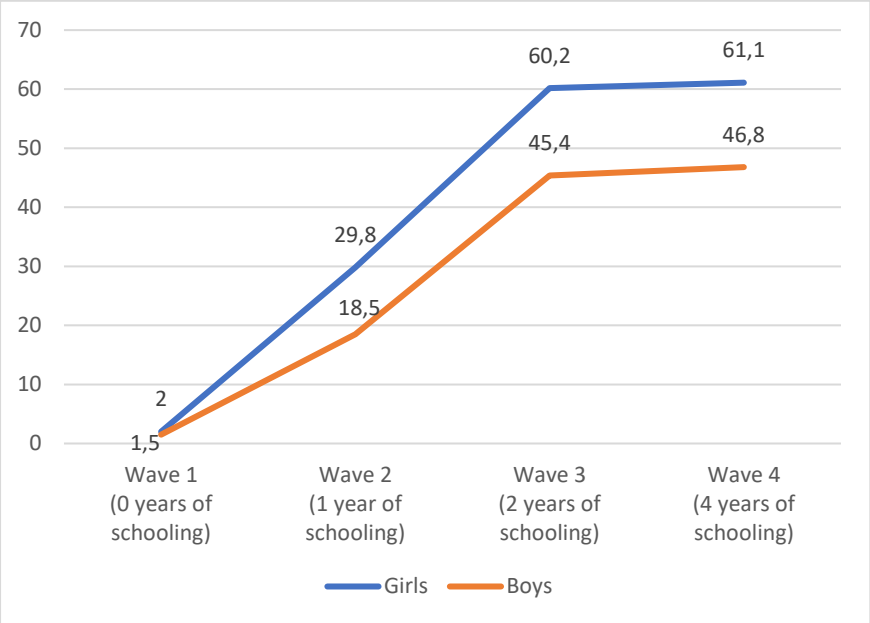


Figure 31: Percentage of learners in each gender group who reached the DBE Grade 1 letter sound benchmark in each wave (excluding repeaters).

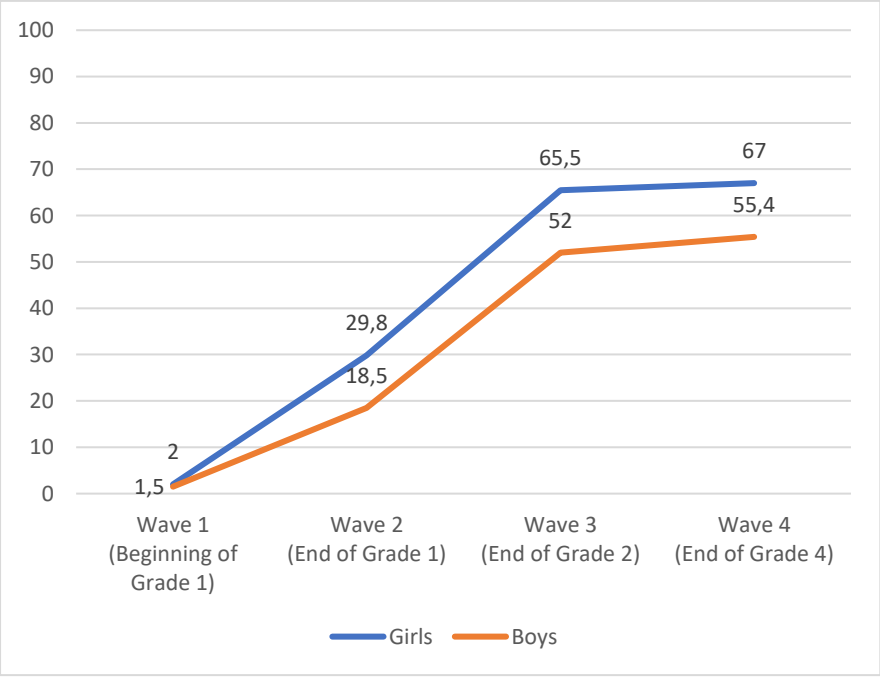


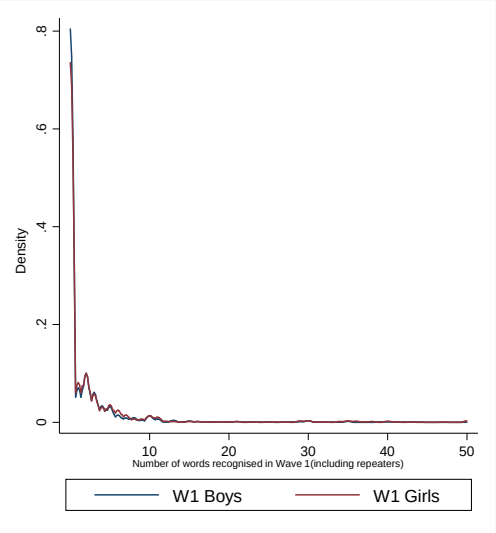
Figure 32 compares the word recognition performance for each gender. At the beginning of Grade 1 (Figure 32a), the shapes of the k-densities are almost identical, although there was a greater percentage of boys who could not recognise a single word and a greater

percentage of girls who recognised between approximately 5 and 8 words. The majority of learners could not recognise any words though, regardless of their gender. After one year of schooling (Figure 32b), there was also a larger percentage of boys who could not recognise a single word, and there was a larger percentage of girls who could read more than 6 words. This gap disappears from approximately 40 words. After two years of schooling (Figure 32c), there was also a larger percentage of boys who could not recognise a single word. From approximately 8 words, the gap between boys and girls narrowed although from approximately 12 it widened again. At 23 words, the percentage of girls and boys was equal, but between 23 and 50 words there was a larger percentage of girls. After four years of schooling (Figure 32d), the difference between girls and boys was quite pronounced. There was a larger percentage of boys who could read between 0 and 30 words and a larger percentage of girls who could read between 30 and 70 words.

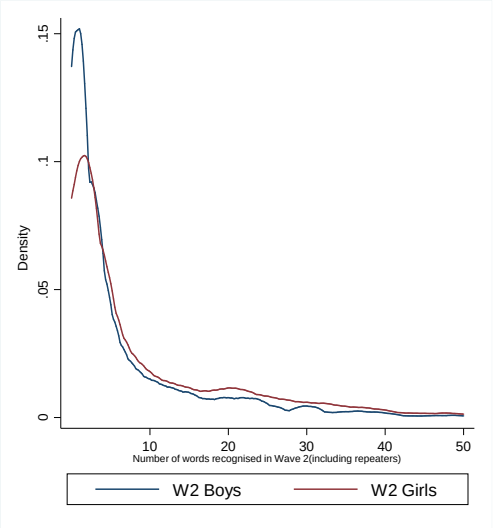
When repeaters are excluded as can be seen in Figure 33, the gaps between the genders were smaller in waves 3 and 4, but especially in wave 4. Again, this makes sense since the majority of repeaters were boys.

Figure 32: K-densities of word recognition performance in the four waves for each gender (including repeaters)

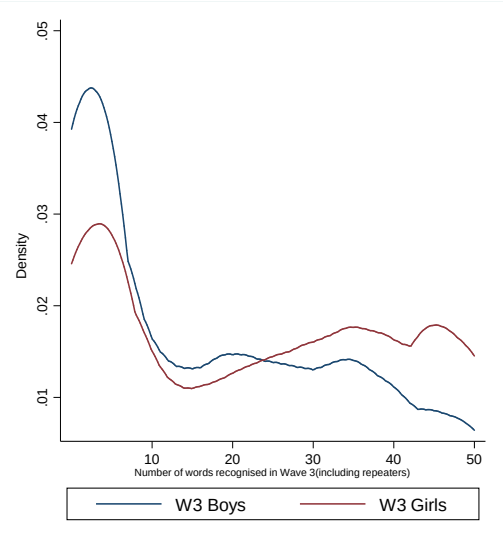
a (0 years of schooling)



b (1 year of schooling)



c (2 years of schooling)



d (4 years of schooling)

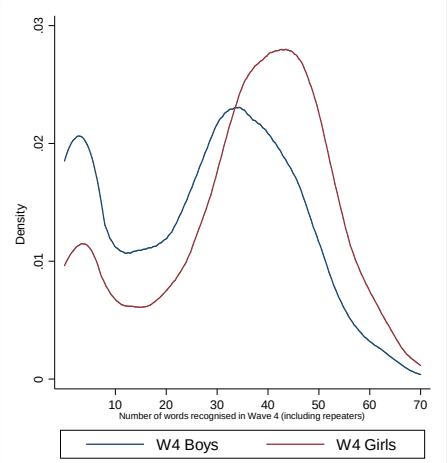


Figure 33: K-densities of word recognition performance in the four waves for each gender (excluding repeaters)

a (End of Grade 2) *b (End of Grade 4)*

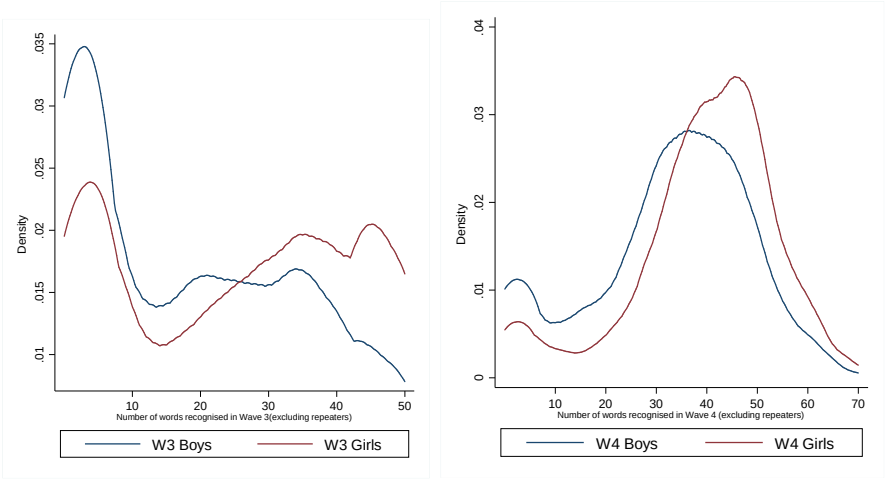


Figure 34 compares the non-word recognition performance for both genders in waves 2 and 3. After one year of schooling (Figure 34a) 69,6 percent of boys could not read a single non-word compared to 61,1 percent of girls. After two years of schooling (Figure 34b), this gap increased as 38,8 percent of boys could not read a single non-word compared to 28,4 percent of girls. There was a larger percentage of boys who read between 0 and 5 non-words, there was no difference between the genders between approximately 5 and 8 non-words but there was a larger percentage of boys again between approximately 8 and 20 non-words. Between approximately 20 and 50 non-words, there was a larger percentage of girls. It is clear that girls read more non-words than boys in both waves. When repeaters are excluded as can be seen in Figure 35, the gap between the two genders diminished slightly. Again, this could be explained by the fact that a greater percentage of repeaters were boys.

Figure 34: K-densities of non-word recognition performance in waves 2 and 3 for each gender (including repeaters)

a (1 year of schooling) *b (2 years of schooling)*

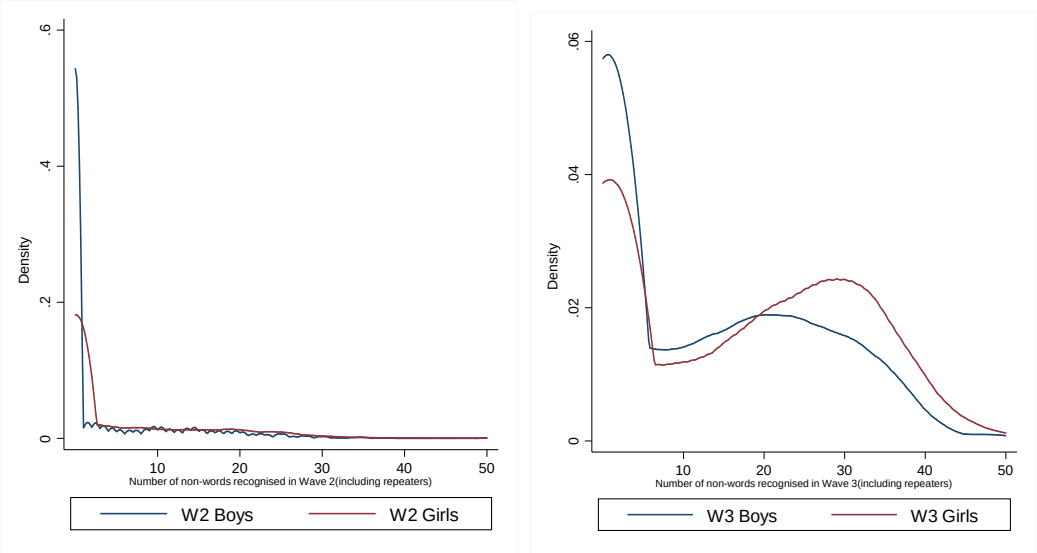


Figure 35: K-densities of non-word recognition performance at the end of Grade 2 for each gender (excluding repeaters)

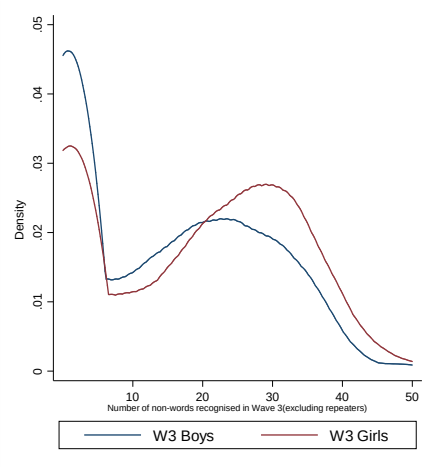


Figure 36 compares the ORF performance for both genders from waves 2 to 4. The red vertical lines represent the Grade 2 and 3 benchmarks (40 WCPM and 60 WCPM respectively). After one year of schooling (Figure 36a), approximately 23 percent of boys were non-readers compared to 9 percent of girls and there was a greater percentage of girls at numerous points between 10 WCPM to 60 WCPM. As can be seen in Figure 36, 7,6 percent of girls and 3,6 percent of boys reached the Grade 2 benchmark and 1,4 percent of girls and 0,4 percent of boys reached the Grade 3 benchmark.

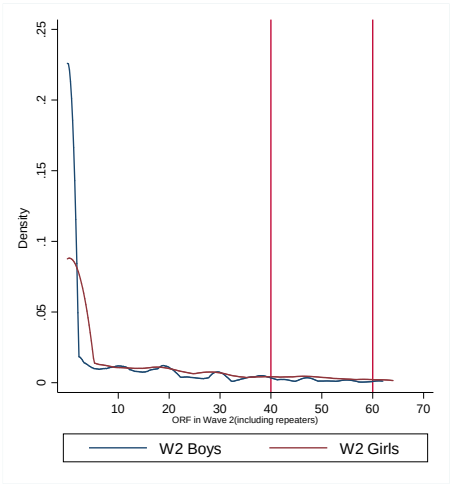
After two years of schooling (Figure 36b), there was also a larger percentage of boys who were non-readers. At approximately 10 WCPM, the gap between the genders disappeared, but between 10 WCPM and the Grade 2 benchmark (40 WCPM), there was a larger percentage of boys. There was a larger percentage of girls who exceeded the benchmarks, however. As can be seen in Figure 37 44,4 percent of girls and 27,3 percent of boys reached the Grade 2 benchmark and 17,4 percent of girls and 7,9 percent of boys reached the Grade 3 benchmark. This means that within the group of girls, there were 1,6 times as many who reached the Grade 2 benchmark than there were in the group of boys and there were more than twice as many within the group of girls who reached the Grade 3 benchmark than there were in the group of boys.

After four years of schooling, learners read two texts. In text 1, as can be seen in Figure 36c, there was a larger percentage of boys who could read between 0 and 50 WCPM. This flipped at the 50 WCPM point. Text 2 (Figure 36d) follows a similar trend although there was a greater percentage of boys up to approximately 55 WCPM (instead of 50 WCPM), and the gap between the genders disappeared after 190 WCPM. Expectedly, a greater percentage of girls reached the benchmarks. For text 1, 73,7 percent of girls and 51,5 percent of boys reached the Grade 2 benchmark and 50,5 percent of girls and 28 percent of boys reached the Grade 3 benchmark. For text 2, 76,2 percent of girls and 58,5 percent of boys reached the Grade 2 benchmark and 59,3 percent of girls and 35,5 percent of boys reached the Grade 3 benchmark. On average, the proportion of girls reaching the Grade 2 benchmark is nearly 40 percent higher than it is for boys and the proportion of girls reaching the Grade 3 benchmark is nearly 75 percent higher than it is for boys.

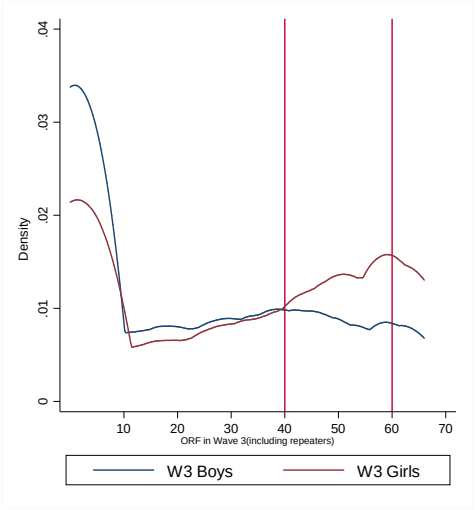
When repeaters are excluded, as can be seen in Figure 38, the gap between the learning profiles was a lot narrower but a greater proportion of girls continued to reach the benchmarks.

Figure 36: K-densities of ORF performance in the waves 2 to 4 for each gender (including repeaters)

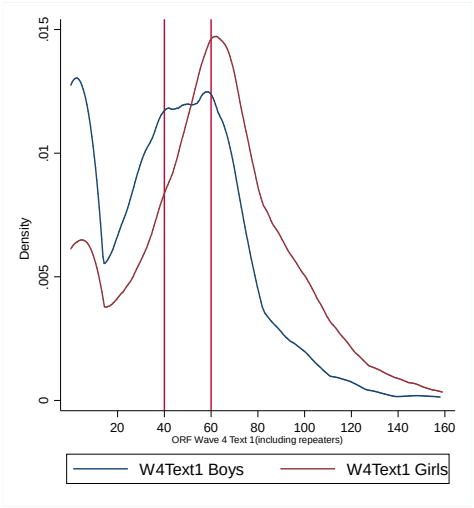
a (1 year of schooling)



b (2 years of schooling)



c (4 years of schooling)



d (4 years of schooling)

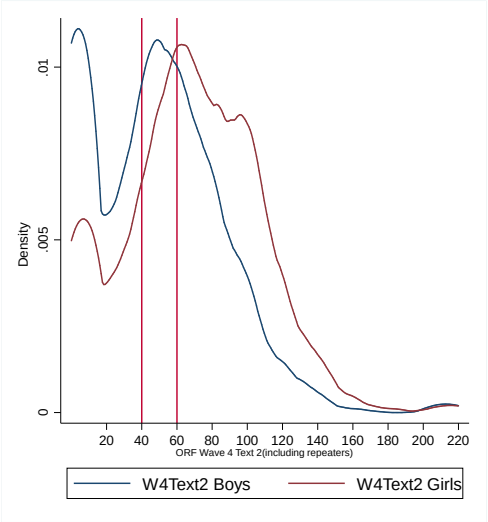


Figure 37a: Percentage of learners in each gender group who reached the DBE Grade 2 benchmark in each wave (including repeaters)

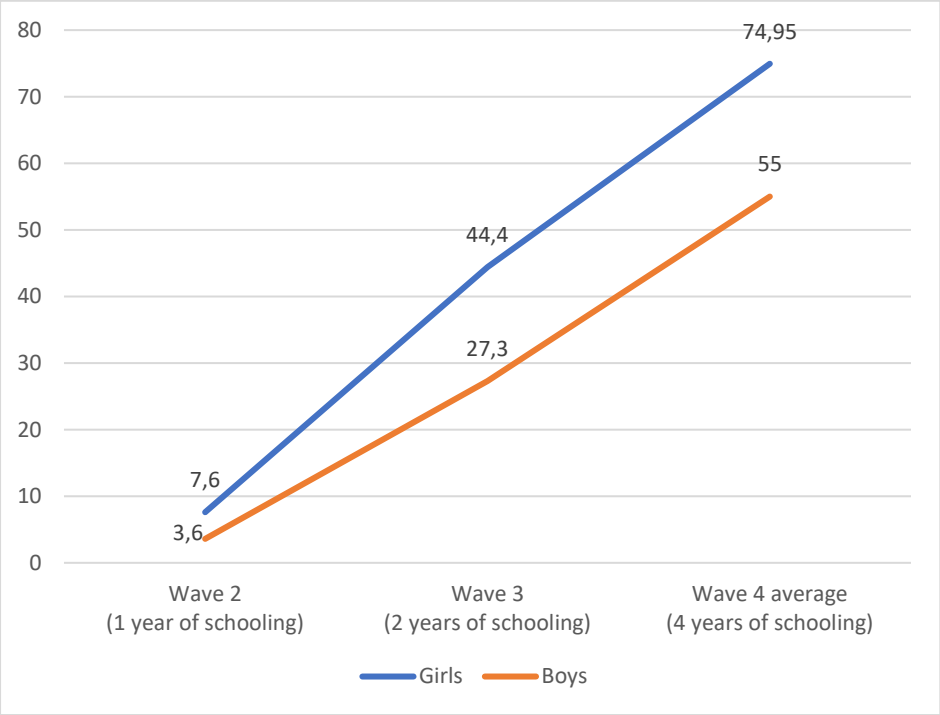


Figure 37b: Percentage of learners in each gender group who reached the DBE Grade 3 benchmark in each wave (including repeaters)

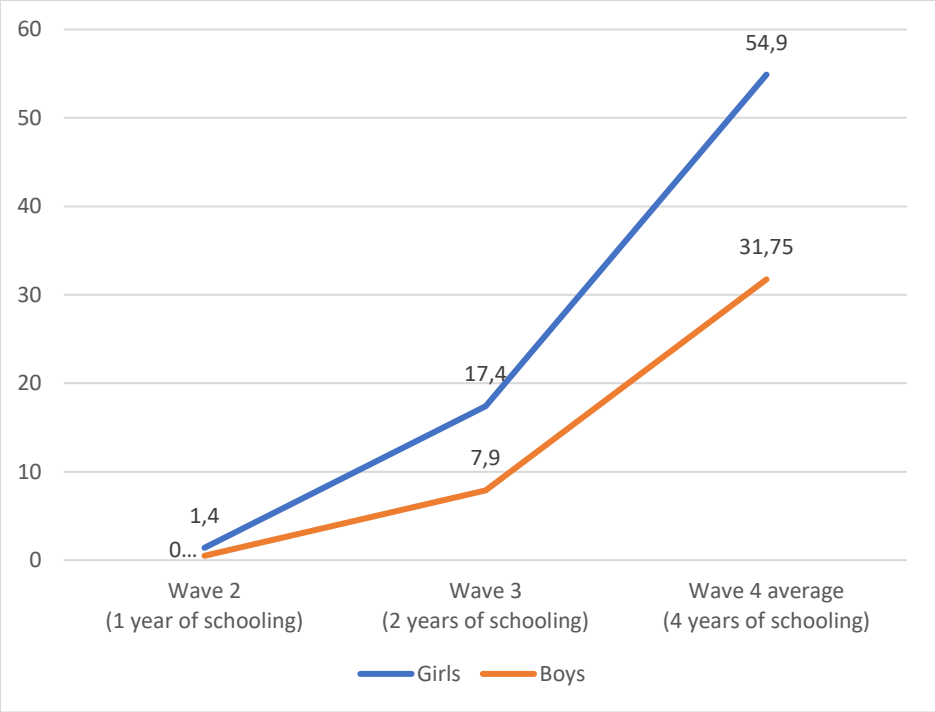


Figure 38a: Percentage of learners in each gender group who reached the DBE Grade 2 benchmark in each wave (excluding repeaters).

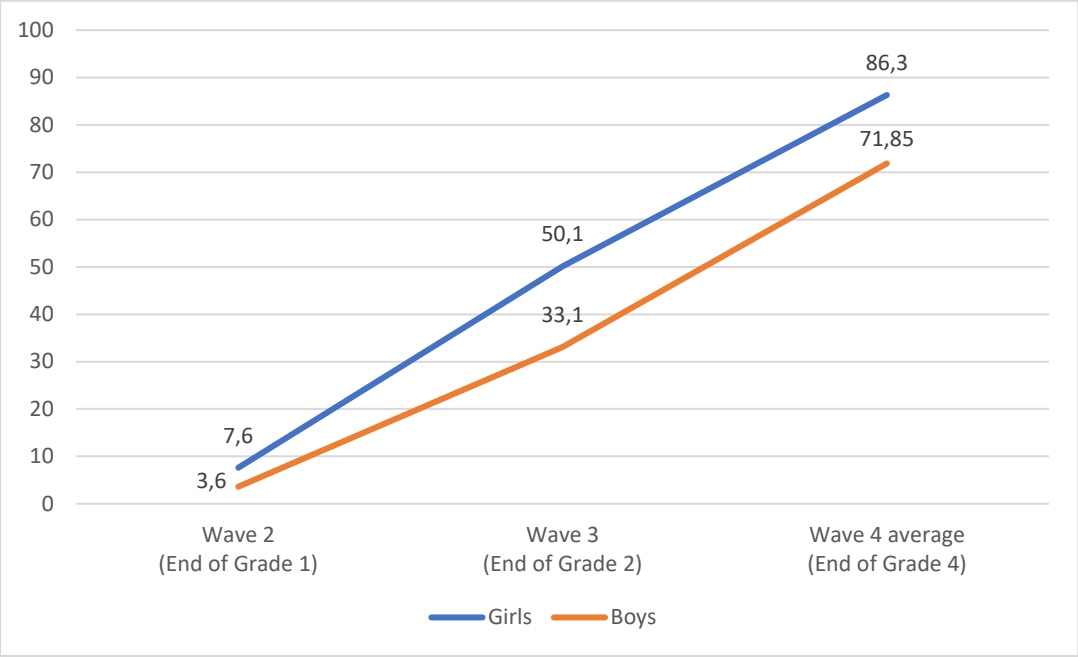
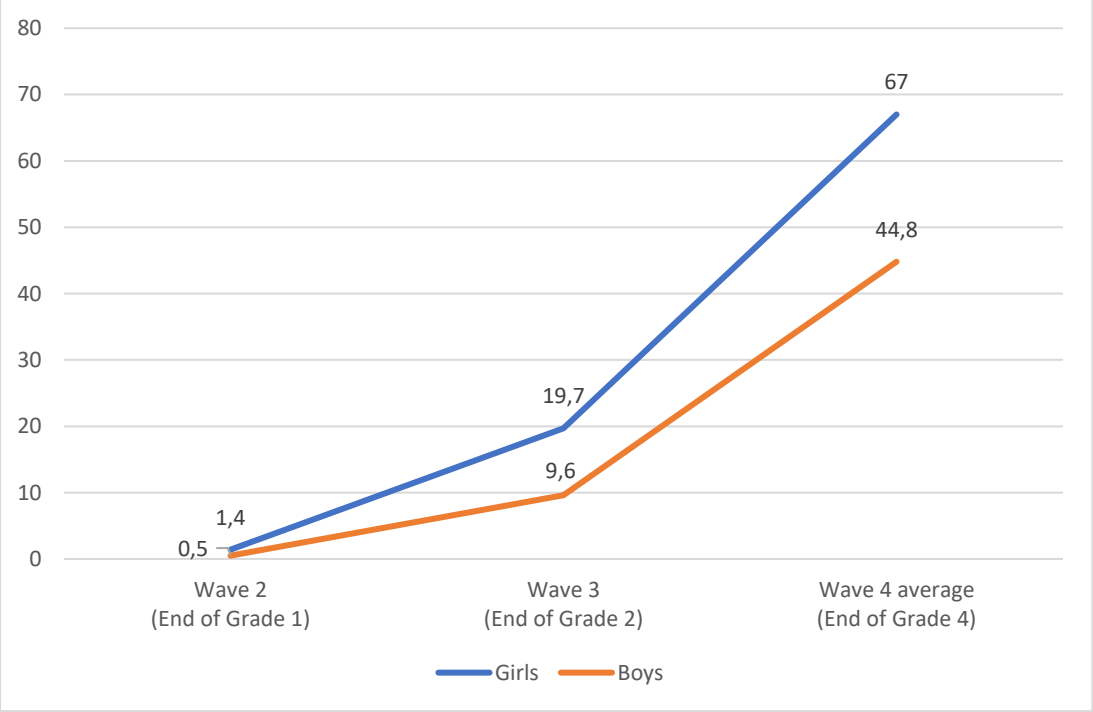


Figure 38b: Percentage of learners in each gender group who reached the DBE Grade 3 benchmark in each wave (excluding repeaters).



Figures 39 and 40 compare the phonological awareness performance for both genders. At the beginning of Grade 1, learners answered 12 questions. They were asked to break up six words into their sound components; provide different words starting with the same two sounds as the three words provided; and provide words ending with the same two sounds as the three words provided. At the end of Grade 1, learners were asked to break up four of the six words asked in wave 1 into their sound components. In Wave 3, learners were asked to identify whether specific sounds were at the beginning, middle or end of two words and they were asked to join two syllables together to form one word.

Figure 39 compares the 4-question phonological awareness scores for both genders in the first two waves. Although a greater proportion of girls achieved higher scores in both waves, the differences are small. It is notable, however, that at the beginning of Grade 1, 62,9 percent of girls could not answer a single question correctly compared to 65,5 percent of boys and at the end of Grade 1 61,5 percent of girls could not answer a single question correctly compared to 72,4 percent of boys. In the previous section, I found that some learners performed worse in the wave 2 test, indicating that they lost phonological awareness skills in Grade 1. By disaggregating these scores by gender, it would appear that the boys were the ones who performed worse in phonological awareness.

In wave 3, as can be seen in Figure 40, the difference between gender performance was small, except for the proportion who achieved a perfect score. 38,6 percent of girls got all the questions correct compared to 32,6 percent of boys.

Figure 39: Percentage of learners within each gender group achieving different scores in the wave 1 and 2 phonological awareness test (including repeaters)

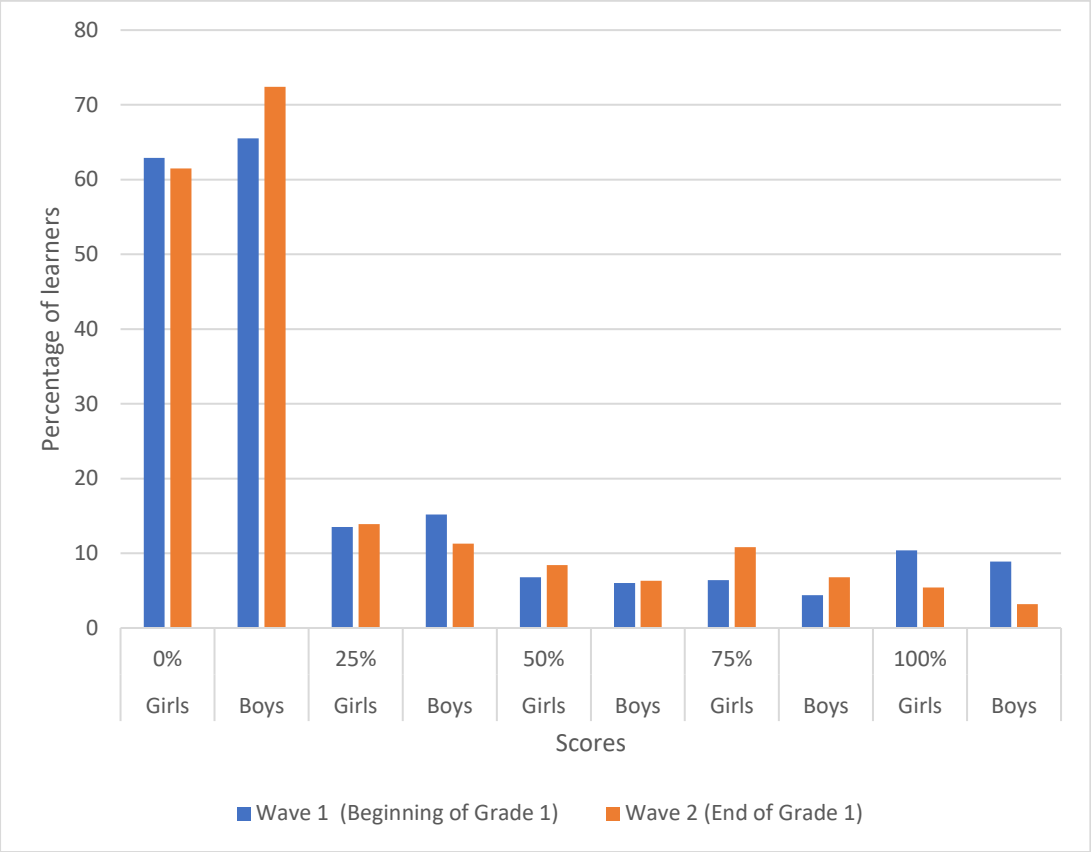
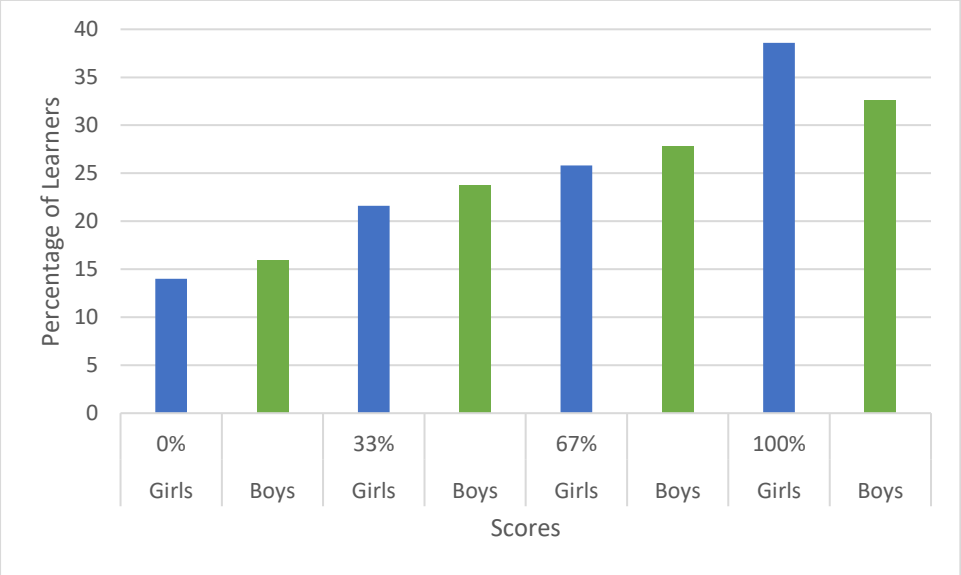


Figure 40: Percentage of learners within each gender group achieving different scores in the phonological awareness test after two years of schooling (including repeaters)



In summary, girls performed better than boys in the wave 4 comprehension test. There was a larger proportion of boys that were alphabetically illiterate and a larger proportion of girls reaching the Grade 1 letter sound benchmarks in every wave (bar the first one where they were approximately equal). While their word recognition performance was similar at the beginning of Grade 1, girls outperformed boys in word recognition after one, two and four years of schooling and in non-words after one and two years of schooling. The results also confirm Wills et al. (2022a) discovery that girls progress faster than boys in ORF. After one, two and four years of schooling there was a greater proportion of boys who were non-readers and after two and four years of schooling, there was a much larger proportion of girls who reached the Grade 2 and 3 benchmarks than there were boys. The phonological awareness results were similar for both genders, however, it is notable that between the beginning and end of Grade 1, boys lost phonological awareness skills whereas girls gained them. Finally, the fact that 62 percent of the repeaters in both waves 3 and 4 were boys is an indication that girls outperformed boys. This is further supported by the finding that 25 percent of the girls who wrote the wave 4 test were in Grade 2 or 3 compared to 36 percent of boys.

4.5 Relationship between Grade 4 comprehension scores and reading components

In the fourth and final section I explored how Grade 4 learners across the distribution of comprehension scores performed in the early grade exercises. This is a backward-looking analysis, and the reading performance in Grade 1 and 2 is conditional on the performance in the Grade 4 comprehension test. As discussed already, in wave 4, learners were given three minutes to read a 159-word story about a dove and an ant. Once the three minutes were up, field workers asked each learner a maximum of eight questions. The number of questions that were asked was dependent on how far the learner read up until in the text. Each of these questions was worth half a point, one point or two points and the maximum number of points a learner could earn was eight.

In order to assess the relationship between reading components and comprehension scores, I performed kernel-weighted local polynomial regressions of every reading component on wave 4 comprehension reading scores. The graphs of the smoothed values that are

displayed below include 95% confidence bands (represented by the grey area). This just means that we can be 95 percent confident that the grey area contains the true mean. (The mean does not have to be on the line). The analysis included repeaters, however, for the purpose of easier reading, “wave 1” will be referred to as “beginning of Grade 1”, “wave 2” will be referred to as “end of Grade 1”, “wave 3” will be referred to as “end of Grade 2” and “wave 4” will be referred to as “end of Grade 4”. It is also important to note that there is no attrition as I restricted the analysis to learners for whom I had wave 4 data. This does mean that some of the Grade 1 and 2 learners that were included in the analysis in sections 4.2 and 4.3 but did not write the comprehension test in Grade 4 are not included in the analysis in this section. This may affect the comparison of this section to the earlier two sections.

Figure 41 displays the relationship between Grade 4 comprehension scores and LCPM in the earlier grades. It is clear that on average, those who achieved a high comprehension score in Grade 4 sounded more letters in the earlier grades than those who had a low comprehension score. At the beginning of Grade 1, none of the learners, irrespective of how well they performed on the Grade 4 comprehension test were able to identify a large number of letters correctly. By the end of Grade 1, however, learners who achieved a comprehension score of 5,5 or above had reached the Grade 1 letter sound benchmark (40 LCPM) whereas those with a score below this had not. This indicates that the first year of schooling is important and those who master their letter sounds are at an advantage. All learners improved by the end of Grade 2, and learners who passed the comprehension test (a score of 4 or above) but did not reach the benchmark by the end of Grade 1 could sound an average exceeding 50 LCPM. No additional learning of letters appeared to take place between the end of Grades 2 and 4. While this could reflect the testing effect (the Grade 4 letter sound test included complex consonant sequences), it could also indicate that there is not a significant improvement in the speed of identifying letters between Grade 2 and 4.

Figure 41: The relationship between Grade 4 comprehension scores and LCPM in the earlier grades

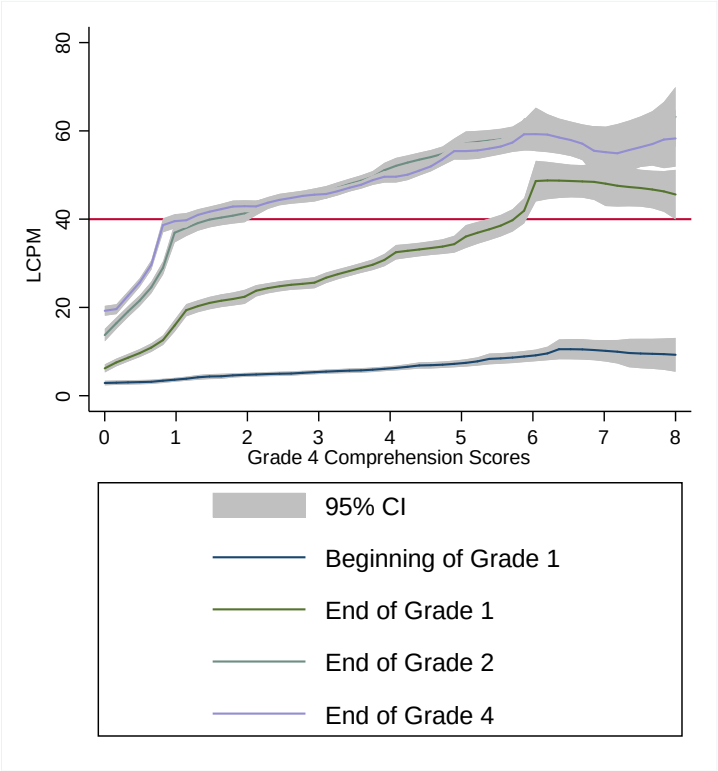


Figure 42 displays the relationship between Grade 4 comprehension scores and word recognition in each grade. Similar to the findings in the previous component, those with high comprehension scores in Grade 4 could recognise more words in the earlier grades than those with lower comprehension scores. At the beginning of Grade 1, none of the learners, irrespective of how they performed on the Grade 4 comprehension test could recognise many words. By the end of Grade 1, however, those who passed the comprehension test could recognise at least 10 words in a list of 50 common words whereas those who failed recognised fewer words. This indicates that Grade 1 is an important year for letter sounds as well as word recognition. By the end of Grade 2, learners who got at least one of the comprehension questions correct made large improvements in word recognition and those who passed could recognise more than 30 words in the 50-word list. Dissimilar to the letter sound finding, learners who got at least one comprehension question correct could recognise more words in a list in Grade 4 than they could in Grade 2, with those who passed able to recognise more than 40 words. This indicates that all learners improved on word recognition between Grades 2 and 4, but it may also reflect the fact that the Grade 4 word recognition word list included more (70) words. Another important point to make is that learners who got 1 question correct could recognise approximately 30 words

by the end of Grade 4 (the point that those who passed the comprehension test reached at the end of Grade 2), but it would seem that this improvement occurred too late for it to translate into high comprehension in Grade 4.

Figure 42: The relationship between Grade 4 comprehension scores and word recognition in the earlier grades

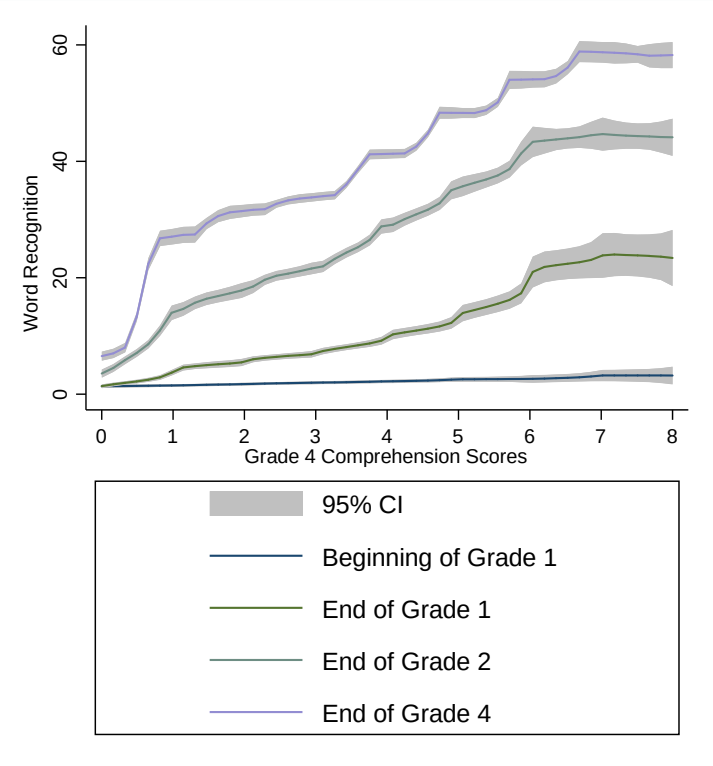


Figure 43 displays the relationship between Grade 4 comprehension scores and non-word recognition. Again, those with high comprehension scores could identify more non-words at the end of Grades 1 and 2 than those with lower comprehension scores. By the end of Grade 1, those who passed the comprehension test could recognise more than 5 non-words but those who failed recognised fewer non-words. By the end of Grade 2, all learners except those who could not get a single comprehension question correct improved, and the learners who passed could recognise more than 20 non-words.

Figure 43: The relationship between Grade 4 comprehension scores and non-word recognition in the earlier grades

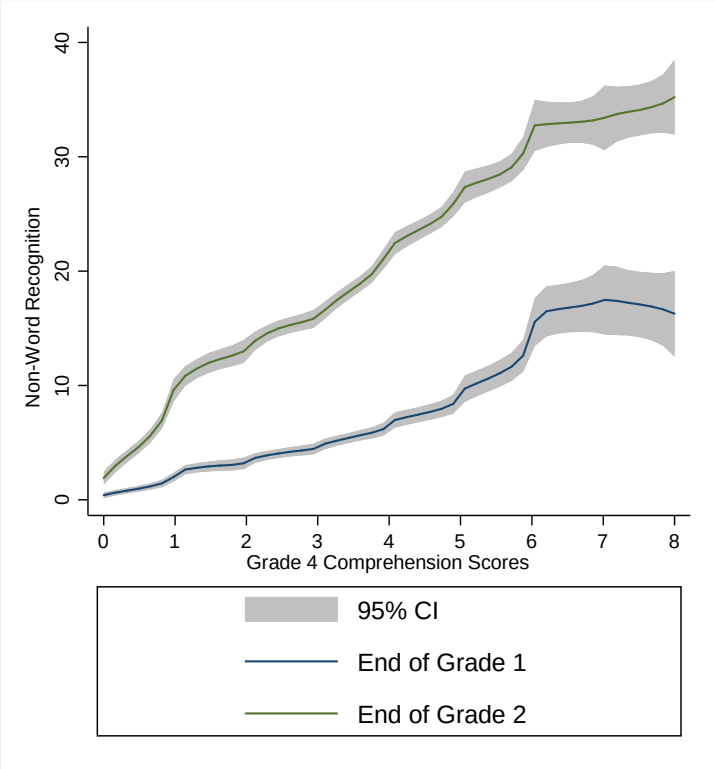


Figure 44 displays the relationship between Grade 4 comprehension scores and ORF at the end of Grades 1, 2 and 4. Once again, those with higher comprehension scores in Grade 4 had a higher ORF in the earlier waves than those with lower comprehension scores. At the end of Grade 1, those with higher comprehension scores could read more than 20 WCPM and some of the learners who achieved a score of 7 already reached the Grade 2 benchmark. By the end of Grade 2, all learners except those who could not comprehend anything they read had improved and those who passed the comprehension test reached the Grade 2 benchmark. It is also interesting to note that those who achieved a comprehension score greater than 6 had reached the Grade 3 benchmark by the end of Grade 2. By the end of Grade 4, all learners, except those who could not get a single comprehension test correct improved their ORF. Moreover, those who passed exceeded the Grade 3 benchmark.

Figure 44: The relationship between Grade 4 comprehension scores and ORF in the earlier grades

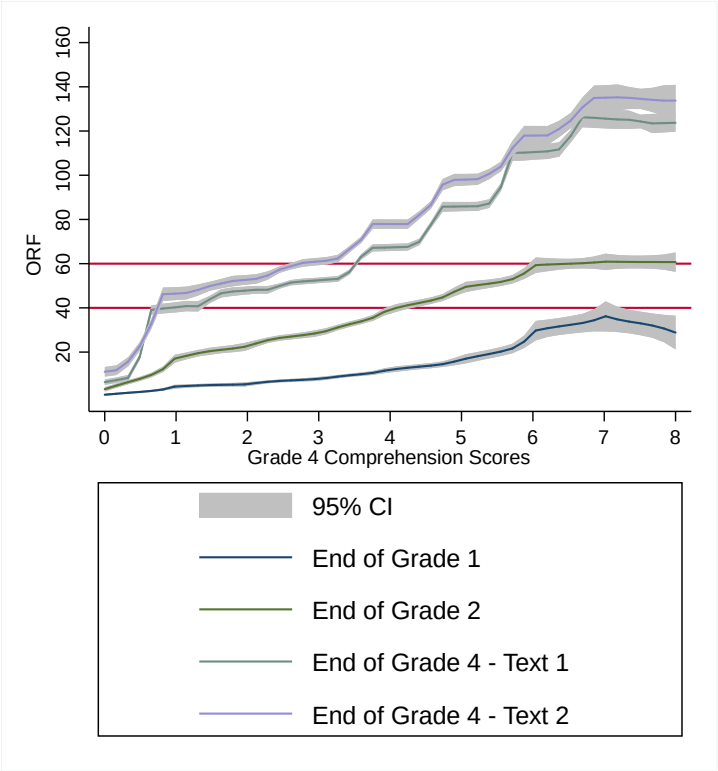
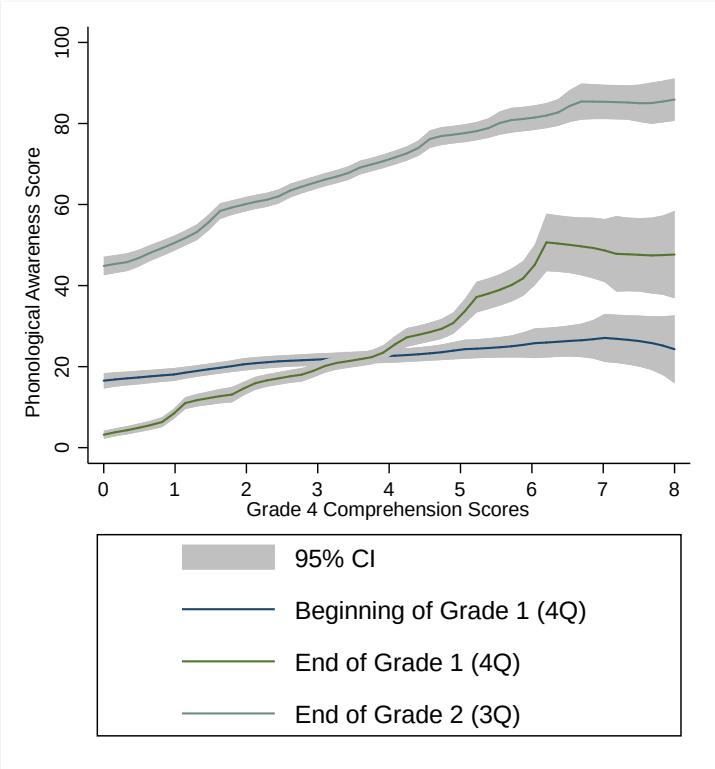


Figure 45 displays the relationship between Grade 4 comprehension scores and phonological awareness scores. At the beginning of Grade 1, all learners, regardless of their comprehension score achieved an average phonological awareness score of approximately 20 percent. By the end of Grade 1, those with a comprehension score that was below three achieved a worse average phonological score whereas those who passed achieved a greater phonological awareness score, the average score being approximately 40. This suggests that it is advantageous for learners to improve their phonological awareness score in the first year of schooling. By the end of Grade 2, learners who passed the comprehension test achieved approximately 75 percent on their phonological awareness test, whereas those who failed achieved a lower score. Since the Grade 2 phonological awareness test was different to the Grade 1 tests, we cannot determine whether phonological awareness skills improved in Grade 2.

Figure 45: The relationship between Grade 4 comprehension scores and phonological awareness scores in the earlier grades



In conclusion, there are clear relationships between Grade 4 comprehension scores and reading components in the early years of schooling. Those with high comprehension scores in Grade 4 had a higher average LCPM and ORF, they recognised more words and non-words and they had greater phonological awareness scores in the earlier grades. They also improved their phonological awareness skills in Grade 1. It is important to remember that 20 percent of Grade 4s (including repeaters) did not get any of the comprehension questions correct, 68 percent achieved a score between half a point and four, and only 12 percent passed. As has already been discovered in the previous sections, the majority of learners did not make enough progress in basic reading components in the early years for comprehension in Grade 4.

In this chapter, there were four sections. In the first section, I explored the iceberg under the surface of the ocean by assessing the performance in reading components after four years of schooling. In the second section, I assessed the pace of progress in reading components over the first four years of schooling. In the third section, I assessed whether girls progressed faster than boys in basic reading components. In the fourth section, I

explored the relationship between comprehension scores in the fourth year of schooling and reading components in the first four years of schooling.

The results in this chapter give us five key insights. The first is that EGRS I learners performed poorly in comprehension after four years of schooling. When repeaters are included, only 12 percent of learners passed the test and one in five learners did not get any of the questions correct. When repeaters are excluded, these figures improved slightly although it is still significant that 83 percent of Grade 4s failed the test and one in ten Grade 4s could not get a single question correct. This largely confirms the PIRLS 2016 finding that 90 percent of learners who wrote the Setswana test could not read for meaning.

The second key insight and the answer to the first research question is that after four years of schooling, a large percentage of learners had also not made adequate progress with basic reading components, namely letter sounds, word recognition and ORF. When repeaters are included, 54 percent of learners did not reach the Grade 1 letter sound benchmark (40 LCPM), 36 percent did not reach the Grade 2 benchmark (40 WCPM) and 57 percent did not reach the Grade 3 benchmark (60 WCPM). Moreover, one in five learners could not recognise more than 10 words in a 70-word reading list. While we may assume that the repeaters were the ones completely responsible for these high percentages, this was not the case. While the repeaters displayed a worse performance, it is significant that two in five Grade 4s did not reach the Grade 1 letter sound benchmark, one in five did not reach the Grade 2 benchmark and at least two in five did not reach the Grade 3 benchmark. Moreover, 6 percent of Grade 4s were non-readers and one in ten could not recognise more than 10 words in a reading list.

The reason many of these learners were ill-prepared in Grade 4 brings us to the third key insight and the answer to the second research question. A large proportion of learners made slow progress in the basic reading components in the first two years of schooling. The majority of learners entered Grade 1 with poor phonological awareness skills, and they did not improve these skills over the next year. Some of the learners (the majority were boys[2])

[2] This finding may be the outcome of a test instrument or attrition.

even got worse. Many learners also entered Grade 1 with a poor letter-sound ability and only half of them made sufficient progress over the next two years. By the end of Grade 1, all learners should be reaching the Grade 1 benchmark, but in the EGRS I sample, only one in five learners reached it. By the end of Grade 2, less than half of the learners had reached it, and this did not change significantly over the next two years. The lack of change indicates that after the first two years of schooling, there was little opportunity for learners to improve their letter sound ability, however, it may also be explained by a testing effect as the wave 4 test included the more difficult complex consonant sequences.

Poor phonological awareness and letter sound skills likely had implications for word recognition, non-word recognition and ORF. By the end of Grade 1, learners should be starting to recognise words but 61 percent of the sample were non-readers, 66 were non-non-word readers and 76 percent could not recognise more than 10 words in a 50-word list of Foundation Phase words. By the end of the second year of schooling, 36 percent of learners were non-readers, 34 percent were non-non-word readers and 41 percent could not recognise more than 10 words. These results indicate that little progress was made with word reading in the first year of schooling, but a lot more progress was made in the second year of schooling. This delay was likely due to the slow progress with letter sounds in the first year of schooling. While progress was made in the second year, it is still important to note that the majority of learners did not reach an adequate level to read with comprehension in Grade 4. By the end of Grade 2 all learners should be reaching the Grade 2 benchmark, but only four in ten Grade 2s achieved this.

The fourth key insight and the answer to the third research question is that girls outperformed boys on three grounds. The first is that 25 percent of girls who wrote the wave 4 test were in Grade 2 or 3 compared to 36 percent of boys. Since boys were more likely to repeat, this suggests that a greater proportion of them were struggling. The second is that 18 percent of girls passed the Grade 4 comprehension test compared to only 7 percent of boys and 13 percent of girls did not get any of the questions correct compared to 25 percent of boys. The third is that girls performed better than boys in every reading component and in every wave and a larger percentage of girls reached the Grade 1, 2 and 3 benchmarks.

The fifth key insight and the answer to the fourth research question is that there was a strong positive relationship between Grade 4 comprehension scores and basic reading components in the earlier years. On average, those who achieved high comprehension scores in their fourth year of schooling sounded more letters, recognised more words and non-words and read more words in a passage of text in a minute in the first four years of schooling compared to those with lower comprehension scores. It is also notable that on average, those who passed the Grade 4 comprehension test improved their phonological awareness skill in Grade 1 whereas those who failed regressed.

CHAPTER FIVE: CONCLUSION, LIMITATIONS AND FUTURE RESEARCH

In South Africa, 78 percent of Grade 4s cannot read for meaning in any language. This research paper has aimed to increase our understanding of this statistic by diving deeper and exploring how quintile 1-3 learners who were taught in Setswana in the North West performed in basic reading components in the first four years of schooling. Specifically, I embarked on a quantitative study and developed True Panel learning profiles using secondary EGRS I data with the intention of answering four research questions:

1. *After four years of schooling, what percentage of learners who were taught in Setswana have acquired basic reading components (letter sounds, word recognition, ORF)?*

2. *What is the pace of progress in basic reading components (letter sounds, word recognition, non-word recognition, ORF and phonological awareness) in the first four years of schooling for learners who are taught in Setswana*

3. *Do girls who are taught in Setswana acquire basic reading skills faster than boys who are taught in Setswana?*

4. *What is the relationship between comprehension scores in the fourth year of schooling and the pace of progress in basic reading components in the earlier years for learners who are taught in Setswana*

In brief, the answer to the first question is that after four years of schooling, children have not made adequate progress with basic reading components, the answer to the second is that progress is slow in the first four years of schooling, but especially in Grade 1, the answer to the third is that girls progress faster than boys and the answer to the fourth is that there is a strong relationship between Grade 4 comprehension scores and basic reading components in the earlier years.

While the EGRS I dataset was not nationally representative, the findings give us a general indication of what the situation is like in the rest of South Africa.

Five general points can be made. The first is that children are entering school with poor phonological awareness and the majority are not improving this skill over the next year. In the last section of the results section, I found that learners who passed the comprehension test in the fourth year of schooling improved their phonological awareness skills in the first year of schooling whereas those who failed the comprehension test did not improve on and even lost the phonological awareness skill. This suggests that more attention needs to be given to phonological awareness in the ECD phase as well as in Grade 1 as it appears to be a fundamental part of learning how to read.

The second point is that the majority of learners enter Grade 1 with poor letter sound skills, many make slow progress with it in the first two years of school and those that have not mastered it by the end of Grade 2, lose the opportunity to do so after Grade 2. By the end of Grade 1, all learners should be reading 40 LCPM but the findings that only one in five learners achieved this by the end of Grade 1 and half of the learners achieved this by the end of Grade 2 suggest that we are still a long way off from achieving this goal.

The third point is that likely due to the poor phonological awareness and letter sound performance, learners make slow progress with word reading. By the end of Grade 1, three in five learners are non-readers (they cannot read a single word in a passage of text) and one in five learners can recognise more than 10 words in a 50-word list. This is an indication that some learners are learning how to read words, but the majority of learners are not. Instead, the majority of learning appears to be taking place in the second year of schooling. By the end of Grade 2, two in five learners were non-readers of a passage of text, three in five could read more than 10 words in a 50-word list and two in five reached the Grade 2 ORF benchmark (40 WCPM). While it is significant that a large percentage of learners cannot read a single word after two years of schooling, there are signs that a group of learners are making progress.

The fourth point is that boys perform worse than girls on three grounds. The first is that there is a larger percentage of boys who are repeaters, the second is that boys perform worse in comprehension and the third is that girls outperform boys in letter sounds, word recognition, non-word recognition, ORF and phonological awareness after at least one year of schooling. While they do appear to enter school with similar skills, these gaps widen as they progress through school. It is also notable that a larger percentage of boys lost phonological awareness skills in Grade 1.

The fifth point is that learners with high comprehension scores in Grade 4 tend to have a history of better phonological awareness, letter sound, word recognition, non-word recognition and ORF scores in the earlier grades than those with lower scores.

This research paper builds on the early, small-scale research (Malda et al., 2014; Wilsenach, 2015 and 2016; Hugo et al., 2005; De Sousa et al., 2011; De Witt & Lessing, 2016; O'Carroll, 2011) as well as the recent large-scale research that has been conducted on early grade reading components in South Africa (Ardington et al., 2020, 2021 and 2022; Mohohlwane et al., 2022; Spaull and Pretorius, 2019; Pretorius and Spaull, 2022; Wills et al., 2022a and 2022b). It also contributes to the literature on the learning gender gap in South Africa (Fleisch & Shindler, 2009; Govender, 2020; Hofmeyr, 2022; Spaull and Makaluza, 2019)

While the findings in this paper are important, it is necessary to note the limitations of EGRS I data. The first is that it is not nationally representative. This means that we cannot make generalisations about the performance in South Africa as a whole or even for non-fee or North West learners. The second is that the phonological awareness tests that were administered in the first three waves were different. The third is that the letter sound and word recognition tests that were administered in wave 4 were more difficult versions of the tests administered in the first three waves. The fourth is that the ORF text was different in every wave. The wave 2 and 3 tests were 95 percent similar, however. The fifth limitation is that non-words were not tested in wave 1 and wave 4, ORF was not tested in wave 1 and phonological awareness was not tested in wave 4. The sixth limitation is that fieldworkers may have made mistakes in the data collection and the DBE may have made mistakes while cleaning the data. The seventh is that the recorded gender of some learners changed across

waves. The eighth is that the grade variables in the first two waves were not clear. My assumption that the wave 2 learners who wrote the test were in Grade 1 might be incorrect. The ninth limitation is attrition. Several learners who participated in the study in wave 1 left the school over the next four years. If these learners were achieving better scores, the results may be underestimates of the true scores. The tenth limitation is that learners were not tested in Grade 3. This means that we do not know how much learning took place in Grade 3 and how much took place in Grade 4. Moreover, a limitation with the kernel-weighted local polynomial regression that was used to explore the relationship between comprehension and reading components is that it does not account for outliers.

There are numerous policy implications of this research. The fact that children enter grade 1 with poor phonological awareness skills suggests a need for early interventions in the ECD phase. Teachers should receive training and support in incorporating activities that enhance phonological awareness such as rhyming, segmenting and blending sounds as this could potentially improve the reading readiness of learners. In addition, the finding that children make slow progress in letter sound recognition during Grade 1 highlights the importance of systematic and targeted instruction in phonics. Up to this point in time, research suggests that in South Africa, structured learning programmes using lesson plans, integrated reading materials and professional support (coaching or Teaching Assistants) have the greatest impact on reading (Ardington, 2021, 2023a, 2023b; Ardington & Henry, 2021; Ardington & Meiring, 2020; Cilliers, 2020; Fleisch & Alsofrom, 2022; Kotze, et al., 2019; Meiring, 2021; Sapire et al., 2022; Taylor & Spaull, 2022; Taylor et al., 2017). Qualitative findings emerging from these studies have also provided useful insight into the reasons this combination of tools works. Specifically, an EGRS II study conducted by Alsofrom (2018) revealed that teachers' apprehension and distrust towards change and external criticism can be mitigated by the presence of a coach who provides helpful suggestions in the form of demonstrations and concrete instructions (in the form of lesson plans) as well as encouragement and support. In addition, Fleisch and Dixon (2019) note the important role that time and space play and how the EGRS intervention resulted in better pacing and sequencing of lessons, better coverage of curriculum content, the imposition of routines and more effective use of classroom space. Another qualitative study emerging from the Funda Wande literacy interventions in the Eastern Cape, Limpopo and the Western Cape draws attention to six

behaviours that are still observed in classrooms and are blockages to reading instruction (Hoadley et al., 2022). These are teachers treating learners as a largely passive, collective and undifferentiated group, teachers lacking a systematic pedagogic orientation to text, teachers holding low expectations of learners (resulting in slow pacing, simplified tasks and collective chorusing), inefficient usage of time (no predictable and stable instructional routines, slow pacing, insufficient time allocated to reading, excessive time allocated to copying and marking), teachers providing learners with little feedback and teachers lacking energy for sustained, purposeful and engaged teaching (Hoadley et al., 2022). These pedagogical issues need to be addressed while keeping the change mechanism process in mind.

The findings in this paper have also indicated that within a single grade, and possibly even within the same classrooms, learners exhibit varying levels of reading proficiency. This has implications for teachers as each of these children may have different instructional needs (read more about this in Stern et al., 2018). It would be helpful to teachers then if they were taught how to measure the reading skills of their learners so that they can tailor appropriate reading instruction for each of them. Of course, this may prove to be a difficult task in the context of large class sizes (Spaull et al., 2022). On the topic of assessment, it would also be valuable if reading skills were measured on a provincial or national level. Since the discontinuation of the ANAs, the only learners who take part in nationally comparable assessments are the matric learners. As has been clearly demonstrated in this paper – and many others – it is too late to intervene in the last year of school, and the necessity for an invaluable instrument for monitoring and accountability is required at a much lower level (possibly Grade 3) (Nuga Deliwe & Van der Berg, 2022).

Finally, these findings call to attention the existence of the literacy gap between girls and boys. Since boys are trailing behind at school, it may be necessary to provide them with remedial instruction. However, on the whole, girls are also performing poorly and so it would likely be more advantageous for boys if learning profiles were steepened for all learners – and the system improved as a whole – as opposed to only focusing on reducing the gender gap (Akmal & Pritchett, 2021; Crouch et al., 2021; Kaffenberger & Pritchett, 2020) .

In the future, it will be advantageous if more large-scale research was conducted on the link between early grade phonological awareness skills and comprehension. Findings in this paper indicated that phonological awareness is potentially a missing puzzle piece in the South African reading crisis, however, due to the various limitations in the EGRS I data, I was unable to explore this in its entirety. It would also be useful to generate the learning profiles of the reading components that were explored in this paper (letter sounds, word recognition, non-word recognition, ORF and phonological awareness) for the other African languages as well as for Afrikaans and EFAL. In this paper I found, for example, that majority of the letter sound and word learning appears to take place in the second grade, after which learners stop learning about letter sounds, and so it would be interesting to assess whether this is also the case in the Nguni languages. It would also be interesting to assess whether the pro-girl reading component gap exists in these other languages. Finally, it should be mentioned that due to Covid, the learning profiles in this paper are likely overestimates of the true situation today. When the PIRLS 2021 results are published in May, we will get a better idea of the true impact of Covid on reading comprehension, however, it would be advantageous to explore the learning profiles of basic reading components for learners who started school right before, during or after Covid.

References

Akmal, M., & Pritchett, L. (2021). Learning equity requires more than equality: Learning goals and achievement gaps between the rich and the poor in five developing countries⋆. International Journal of Educational Development, 82, 102350.

Alsofrom, K. (2018). The Second Early Grade Reading Study How and why does coaching work to improve teaching practices in the EGRS II? An examination of causal mechanisms

Ardington, C. (2021). Impact evaluation of Funda Wande Coaching intervention Midline Findings - Midline II

Ardington, C. (2023a). Funda Wande Eastern Cape Workbooks Evaluation

Ardington, C. (2023b). Funda Wande Limpopo Evaluation - Second Midline Report

Ardington, C., & Henry, J. (2021). Funda Wande Limpopo Evaluation - Midline I Report

Ardington, C., & Meiring, T. (2020). Impact evaluation of Funda Wande coaching intervention midline findings. Cape Town: SALDRU, University of Cape Town. Available at: https://fundawande. org/news/impact-evaluation-of-funda-wande-coaching-intervention-midlinefindings-23 (accessed 7 July 2020).

Ardington, C., Mohohlwane, N., & Barends, Z. (2022). Benchmarks Report: Afrikaans Early Grade Reading. DBE.

Ardington, C., Wills, G., Pretorius, E., Deghaye, N., Menendez, A., Mohohlwane, N., Mtsatse, N., & Van der Berg, S. (2020). Technical Report: Benchmarking early grade reading skills in Nguni languages. ReSEP, Stellenbosch University. Cape Town: SALDRU, University of Cape Town. Chicago: NORC at the University of Chicago.

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

Avery, M. (2013). Literature review for local polynomial regression. Unpublished manuscript.

Bartlett, L., Dowd, A. J., & Jonason, C. (2015). Problematizing early grade reading: Should the post-2015 agenda treasure what is measured?. International Journal of Educational Development, 40, 308-314.

Beatty, A., Berkhout, E., Bima, L., Pradhan, M., & Suryadarma, D. (2021). Schooling progress, learning reversal: Indonesia's learning profiles between 2000 and 2014. International Journal of Educational Development, 85

Below, J. L., Skinner, C. H., Fearrington, J. Y., & Sorrell, C. A. (2010). Gender differences in early literacy: Analysis of kindergarten through fifth-grade dynamic indicators of basic early literacy skills probes. School Psychology Review, 39(2), 240-257.

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

Crouch, L., Kaffenberger, M., & Savage, L. (2021). Using learning profiles to inform education priorities: An editors' overview of the Special Issue. International Journal of Educational Development, 86, 102477.

De Sousa, D. S., Greenop, K., & Fry, J. (2010). The effects of phonological awareness of Zulu-speaking children learning to spell in English: A study of cross-language transfer. British Journal of Educational Psychology, 80(4), 517-533.

De Witt, M., & Lessing, A. (2016). The influence of a school readiness program on the language and phonological awareness skills of preschool children in rural areas of South Africa. Australasian Journal of Early Childhood, 41(1), 106-114.

Dubeck, M. M., & Gove, A. (2015). The early grade reading assessment (EGRA): Its theoretical foundation, purpose, and limitations. International Journal of Educational Development, 40, 315-322.

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

Fleisch, B., & Dixon, K. (2019). Identifying mechanisms of change in the early grade reading study in South Africa. South African Journal of Education, 39(3).

Fleisch, B., & Shindler, J. (2009). Gender repetition: school access, transitions and equity in the 'Birth-to-Twenty'cohort panel study in urban South Africa. Comparative Education, 45(2), 265-279.

GEM. (2015). The Southern and Eastern Africa consortium for monitoring educational quality. Assessment Gems Series

Gove, A., & Cvelich, P. (2011). Early reading: Igniting education for all. A report by the early grade learning community of practice. RTI International.

Govender, R. (2020) Gender differences in Foundation Phase learners' literacy achievements. Journal of Research & Method in Education, 10(3)

Hoadley, U., Boyd, C., Sandan, A., & Njara, Q. (2022). Funda Wande (literacy) qualitative study: 2021 implementation

Hofmeyr, H. (2022). Why do girls do better? Unpacking South Africa's gender gap in PIRLS and TIMSS. International Journal of Educational Development, 94, 102648.

Howie, S.J., Combrinck, C., Roux, K., Tshele, M., Mokoena, G.M., & McLeod Palane, N. (2017). PIRLS Literacy 2016 Progress in International Reading Literacy Study 2016: South African Children's Reading Literacy Achievement. Pretoria: Centre for Evaluation and Assessment. https://www.up.ac.za/media/shared/164/ZP_Files/pirls-literacy-2016_grade-4_15-dec-2017_low-quality.zp137684.pdf

Hugo, A. J., Le Roux, S. G., Muller, H., & Nel, N. M. (2005). Phonological awareness and the minimising of reading problems: A South African perspective. Journal for Language Teaching= Ijenali Yekufundzisa Lulwimi= Tydskrif vir Taalonderrig, 39(2), 210-225.

Kaffenberger, M. (2019). A Typology of Learning Profiles: Tools for Analysing the Dynamics of Learning. RISE Insight Series

Kaffenberger, M., & Pritchett, L. (2020). Aiming higher: Learning profiles and gender equality in 10 low- and middle-income countries. International Journal of Educational Development, 79

Korpipää, H., Koponen, T., Aro, M., Tolvanen, A., Aunola, K., Poikkeus, A., Lerkkanen, M., & Nurmi, J. (2017). Covariation between reading and arithmetic skills from Grade 1 to Grade 7. Contemporary Educational Psychology, 51, 131-140

Kotze, J., Fleisch, B., & Taylor, S. (2019). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. International Journal of Educational Development, 66, 203-213.

Malda, M., Nel, C., & Van de Vijver, F. J. (2014). The road to reading for South African learners: The role of orthographic depth. Learning and Individual Differences, 30, 34-45.

Meiring, C. (2021). Impact evaluation of Funda Wande in-service teacher coaching intervention: Findings from the first year (Master's thesis, Faculty of Commerce).

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P (2007). PIRLS 2006 International Results in Reading. Chestnut Hill, MA: TIMSS and PIRLS International Study Centre: Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). PIRLS 2011 International Results in Reading. Chestnut Hill, MA: TIMSS and PIRLS International Study Centre: Boston College.

Mullis, I. V. S, Martin, M. O, Foy, P. & Hooper, M. (2017). PIRLS 2016 International Results in Reading. Chestnut Hill, MA: TIMSS and PIRLS International Study Centre: Boston College.

Mohohlwane, N., Ardington, C., & Wills, G. (2022). A review of recent efforts to benchmark early reading skills in South African languages. In Spaull, N. & Pretorius, E. (Eds) Early Grade Reading in South Africa. Cape Town: Oxford University Press

NCES. (2016). Progress in International Reading Literacy Study (PIRLS) - Overview. https://nces.ed.gov/surveys/pirls/

Nuga Deliwe, C., & Van der Berg, S. (2022). The role of assessment in Foundation Phase improvement: The Annual National Assessments and beyond. In Venkat, H. & Roberts, N. (Eds) Early Grade Mathematics in South Africa. Cape Town: Oxford University Press.

O'Carroll, S. (2011). An exploratory study of early letter-sound knowledge in a low socio-economic context in South Africa. Reading & Writing, 2(1), 7-26.

Pretorius, E., & Spaull, N. (2022). A New Era of Reading Research in South Africa Coming of Age and Accounting for Empirical Regularities. In Spaull, N. & Pretorius, E. (Eds) Early Grade Reading in South Africa. Cape Town: Oxford University Press.

Pritchett, L., & Sandefur, J. (2020). Girls' schooling and women's literacy: schooling targets alone won't reach learning goals. International Journal of Educational Development, 78, 102242.

Saito, M. (2011). Trends in the magnitude and direction of gender differences in learning outcomes. SACMEQ: UNESCO International Institute for Educational Planning, 70.

Sapire, I., Isaac, P., Mpofu, S., Sako, L., Seoloana, M., Ndamase, T, & Mafilika, V. (2022). The implementation of the Bala Wande programme in Grade 1 in three provinces: lessons learned. In Spaull, N. & Taylor, S. (Eds.), Early Grade Reading and Mathematics Interventions in South Africa. Cape Town: Oxford University Press.

Schaefer, M., & Kotzé, J. (2019). Early reading skills related to Grade 1 English Second Language literacy in rural South African schools. South African Journal of Childhood Education, 9(1), 1-13.

Sigmundsson, H., Eriksen, A. D., Ofteland, G. S., & Haga, M. (2017). Sound knowledge: Exploring gender differences in children when they start school regarding knowledge of large letters, small letters, sound large letters, and sound small letters. Frontiers in Psychology, 8, 1539.

Silberstein, J. (2021). Measuring, Visualising, and Simulating Solutions to the Learning Crisis: New Evidence from Learning Profiles in 18 Countries. RISE Insight Series

Smith, C., Stone, R., & Comings, J. (2012). Field study report: Literacy policy and practice in Ethiopia: Building on the TELL program and EGRA results. Amherst, MA: Centre for International Education, University of Massachusetts, Amherst.

Sørensen, T. B. (2015). Review of Early Grade Reading Assessment (EGRA).

Spaull, N., Courtney, P. & Qvist, J. (2022). Mathematical stunting in South Africa: An analysis of Grade 5 mathematics outcomes in TIMSS 2015 and 2019. In Venkat, H. & Roberts, N. (Eds), Early Grade Mathematics in South Africa. Cape Town: Oxford University Press.

Spaull, N., & Kotze, J. (2015). Starting Behind and Staying Behind in South Africa: The case of insurmountable learning deficits in mathematics. International Journal of Educational Development, 41

Spaull, N., & Makaluza, N. (2019). Girls Do Better: The pro-female gender gap in learning outcomes in South Africa 1995–2018. Agenda, 33(4), 11-28.

Spaull, N., & Pretorius, E. (2019). Still Falling at the First Hurdle: Examining Early Grade Reading in South Africa. In N. Spaull & J. D Jansen (Eds.), South African Schooling: The Enigma of Inequality. (147-168). Springer

Spaull, N. & Pretorius, E. (2022). Early Grade Reading in South Africa. Cape Town: Oxford University Press

Spaull, N., Pretorius, E., & Mohohlwane, N. (2020). Investigating the comprehension iceberg: Developing empirical benchmarks for early-grade reading in agglutinating African languages. South African Journal of Childhood Education, 10(1), 1-14.

Spaull, N. & Taylor, S. (2022). Early Grade Reading and Mathematics Interventions in South Africa. Cape Town: Oxford University Press

Spaull, N. (2023). 2023 Reading Panel Background Report. 2030 Reading Panel.

Stern, J. M., Dubeck, M. M., & Dick, A. (2018). Using Early Grade Reading Assessment (EGRA) data for targeted instructional support: Learning profiles and instructional needs in Indonesia. International Journal of Educational Development, 61, 64-71.

Taylor, S., Cilliers, J., Prinsloo, C., Fleisch, B., & Reddy, V. (2017). The early grade reading study: Impact evaluation after two years of interventions. Pretoria: Department of Basic Education. Accessed August, 16, 2022.

Taylor, S. & Spaull, N. (2022). What works and what scales? Returning to a tradition of evidence-based system-wide programmes. In Spaull, N. & Taylor, S. (Eds), Early Grade Reading and Mathematics Interventions in South Africa. Cape Town: Oxford University Press

UNESCO (2023). What you need to know about literacy. Available at https://www.unesco.org/en/literacy/need-know

UNICEF. (2022, June). Primary School Age Education - UNICEF DATA. UNICEF DATA. https://data.unicef.org/topic/education/primary-education/

Van Staden, S., & Gustafsson, M. (2022). What a decade of PIRLS results reveals about early grade reading in South Africa: 2006, 2011, 2016. In Spaull, N. & Pretorius, E. (Eds) Early Grade Reading in South Africa. Cape Town: Oxford University Press.

Wills, G., Ardington, C., & Lesang Sebaeng, M. (2022a). What do we know about foundational skills in early grade home language reading in South Africa? Empirical evidence from EGRA studies 2015-2021. In Spaull, N. & Pretorius, E. (Eds) Early Grade Reading in South Africa. Cape Town: Oxford University Press.

Wills, G., Ardington, C., Pretorius, E. J., Pooe, E., & Ramagoshi, R. (2022b). Benchmarking early grade reading skills: Setswana and English First Additional Language. Technical report. Khulisa Management Services; Pretoria: Department of Basic Education.

Wilsenach, C. (2015). Receptive vocabulary and early literacy skills in emergent bilingual Northern Sotho-English children. Reading & Writing-Journal of the Reading Association of South Africa, 6(1), 1-11.

Wilsenach, C. (2016). Identifying phonological processing deficits in Northern Sotho-speaking children: The use of non-word repetition as a language assessment tool in the South African context. South African Journal of Communication Disorders, 63(2), 1-11.

Wilsenach, C., & Makaure, P. (2018). Gender effects on phonological processing and reading development in Northern Sotho children learning to read in English: A case study of Grade 3 learners. South African Journal of Childhood Education, 8(1), 1-12.

Appendices

Appendix A: Ethics Clearance Certificate

SCHOOL OF EDUCATION ETHICS COMMITTEE

CONSTITUTED UNDER THE UNIVERSITY HUMAN RESEARCH ETHICS COMMITTEE (NON-MEDICAL)

CLEARANCE CERTIFICATE	**PROTOCOL NUMBER:** 2022ECE168M
PROJECT TITLE	Learning profiles and reading benchmarks: Trends from South African data on Early Grade reading
INVESTIGATOR	Qvist Jessica
SCHOOL/DEPARTMENT OF INVESTIGATOR	**WSOE**
DATE CONSIDERED	21 October 2022
DECISION OF THE COMMITTEE	Approved unconditionally
RISK LEVEL	No Risk
EXPIRY DATE	Date of submission of the Research Report
ISSUE DATE OF CERTIFICATE	**CHAIRPERSON** Dr. Batseba Mofolo-Mbokane

cc: Prof. Brahm Fleish

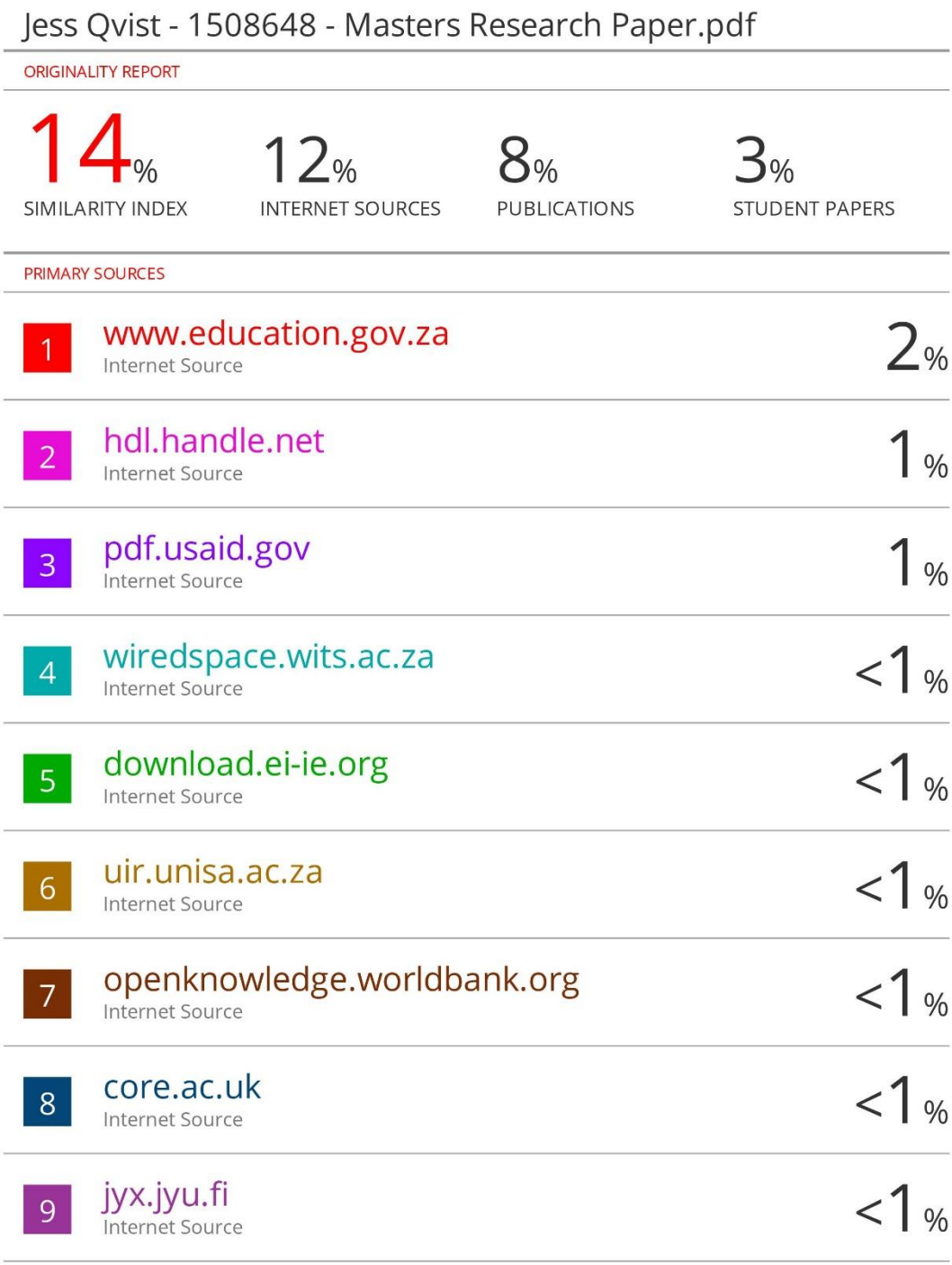
DECLARATION OF INVESTIGATOR

To be completed in duplicate and **ONE COPY** returned to the Chairperson of the School/Department ethics committee.

I fully understand the conditions under which I am are authorized to carry out the abovementioned research and I guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee.

Signature _____ **Date** 25 / 10 / 22

Appendix B: Turnitin Report



Jess Qvist - 1508648 - Masters Research Paper.pdf

ORIGINALITY REPORT

14%	12%	8%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.education.gov.za Internet Source	2%
2	hdl.handle.net Internet Source	1%
3	pdf.usaid.gov Internet Source	1%
4	wiredspace.wits.ac.za Internet Source	<1%
5	download.ei-ie.org Internet Source	<1%
6	uir.unisa.ac.za Internet Source	<1%
7	openknowledge.worldbank.org Internet Source	<1%
8	core.ac.uk Internet Source	<1%
9	jyx.jyu.fi Internet Source	<1%