SUMMARY REPORT:

Benchmarking early grade reading skills in Nguni languages

Cally Ardington, Gabrielle Wills, Elizabeth Pretorius, Nicola Deghaye, Nompumelelo Mohohlwane,

Alicia Menendez, Nangamso Mtsatse, and Servaas van der Berg

OCTOBER 2020

basic education
Department:
Basic Education
REPUBLIC OF SOUTH AFRICA

RESEP
Research on Socio-Economic Policy

UNIVERSITEIT
iYUNIVESITHI
STELLENBOSCH
UNIVERSITY

100
1918·2018

NORC
at the
University of
Chicago

SALDRU
Southern Africa Labour and
Development Research Unit

SUMMARY REPORT:

Benchmarking early grade reading skills in Nguni languages

October 2020

Cally Ardington[1], Gabrielle Wills[2], Elizabeth Pretorius[3], Nicola Deghaye[4], Nompumelelo Mohohlwane[5],

Alicia Menendez[6], Nangamso Mtsatse[7], and Servaas van der Berg[8]

1 SALDRU, University of Cape Town and NORC at the University of Chicago.
2 Research on Socioeconomic Policy (ReSEP), Department of Economics, Stellenbosch University.
3 Research Fellow & Emeritus Professor: Department of Linguistics and Modern Languages, Unisa
4 ReSEP, Department of Economics, Stellenbosch University.
5 Deputy Director of Research Coordination, Monitoring and Evaluation, Department of Basic Education, South Africa.
6 Harris School of Public Policy at the University of Chicago & NORC at the University of Chicago
7 Funda Wande and Stellenbosch University, ReSEP.
8 Department of Economics, Stellenbosch University; ReSEP, and National Research Chair in the Economics of Social Policy.

Abstract

While reading for meaning is the goal of reading, many foundational skills need to be mastered before children can read and understand a text on their own. This report establishes thresholds and benchmarks for foundational early grade reading skills that are necessary (albeit not sufficient) to read for meaning in Nguni languages in the South African context. Our approach to establishing benchmarks integrates theoretical understandings of reading development with an exploratory analysis of early grade reading assessment (EGRA) data in three Nguni languages: isiZulu, isiXhosa and siSwati. The data used is the largest available source of information on early grade reading in these languages, with multiple and comparative assessment data points for nearly 16,400 unique learners in the early grades. By rigorously analysing empirical regularities and reading trajectories across these data, we identify the following context-sensitive thresholds and benchmarks: a letter-sounds benchmark at the end of grade 1 of 40 letter-sounds per minute; a lower fluency threshold at the end of grade 2 of 20 words per minute and a fluency benchmark at the end of grade 3 of 35 words per minute.

About this report

This report summarises the key findings and approach used in identifying early grade reading benchmarks and thresholds in three Nguni languages: isiZulu, isiXhosa and siSwati. We encourage the reader of this summary report to refer to the more detailed technical report on which this is based. The technical report is available at www.resep.sun.ac.za.

Foreword

The progress in supporting reading practices through the development and publication of the National Framework for the Teaching of Reading in African Languages in the Foundation Phase as well as the publishing of the Reading Benchmarks Design Report for the Foundation Phase mark two significant landmark contributions to supporting the teaching, assessment and monitoring of reading in the early grades particularly for African languages.

This report takes forward one of the proposed methods for establishing reading benchmarks in the early grades through the re-analysis of evaluation data from reading intervention research implemented across South Africa since 2015.

The proposed Nguni benchmarks put forth in this report bring together the theoretical underpinnings of linguistic relationships and engages with how similar or different these may be with proposals for a sector response. Although there may be data limitations, this marks the largest and most rigorous approach to Nguni language benchmarks to date. Future data collected for each language may be used to challenge to affirm these, however, this report provides a solid foundation for this debate and evidence base.

This work has involved collaborative efforts across a multi-disciplinary team of researchers from various universities and organisations. The work has largely been produced by South Africans with funding support from both local and international organisations. It draws on data collected from multiple early grade reading studies conducted over the past 4 years in South Africa, including the second early grade reading study (EGRS II) undertaken by the Department of Basic Education (DBE). I would like to commend all the researchers within the DBE, universities, Non-Government Organisations (NGOs) and donors who have prioritised this work. I especially appreciate the efforts of the Research in Socio-Economic Policy (ReSEP) group at the University of Stellenbosch who have led the efforts for this report.

I also look forward to the use of these proposed Nguni benchmarks in informing classroom practice, curriculum practice and systemic assessment.

Finally, I look forward to similar reports produced across the South African languages and phases from a range of researchers, NGOS' and funders.

MR HM MWELI

DIRECTOR-GENERAL

Acknowledgments

Contents

Figures

Tables

Abbreviations

Cpm	Correct per minute
Cwpm	Correct words read per minute
DBE	Department of Basic Education
DFID	Department for International Development
EGRA	Early Grade Reading Assessment
EGRS II	Early Grade Reading Study II
ESRC	Economic Sciences Research Council
FW	Funda Wande
LFL	Leadership for Literacy
LOLT	Language of Learning and Teaching
LS	Letter-sounds
LSPM	Letter-sounds per minute
ORF	Oral Reading Fluency
PIRLS	Progress in International Reading Literacy Study
RC	Reading comprehension
SPS	Story Powered Schools
SVR	Simple view of reading
USAID	United States Agency for International Development
Zenlit	Zenex Foundation Literacy Project
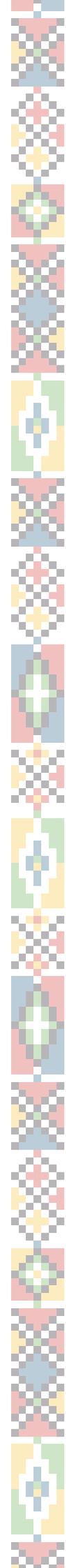
Executive summary

In early 2019 South African president Cyril Ramaphosa articulated a new and clear expectation for basic education: *every child should be able to read for meaning by age 10* (South African Government, 2019). While reading for meaning is the goal of reading, reading is a complex and hierarchical process. A range of foundational reading subskills need to be mastered before one can comprehend or understand what is in a text. For example, knowledge is required of the 'code' of the language in which learners are reading, which we refer to as decoding skills. This report establishes thresholds and benchmarks for some of these foundational early grade decoding skills in three Nguni languages: isiZulu, isiXhosa and siSwati.

Why we need reading benchmarks

The most recent Progress in International Reading Literacy Study (PIRLS 2016) showed that 78% of grade 4 learners, mostly tested in their home language, could not reach the low international benchmark – an indicator of the ability to read for meaning (Howie et al. 2017). Reading comprehension assessments such as PIRLS can identify processes of reading comprehension that learners have not mastered. However, they cannot specify which foundational aspects of reading pose problems for learners who struggle to understand what they are reading. To identify foundational decoding problems, early grade assessments of reading are required. In this report, we use early grade assessment data to establish benchmarks and thresholds.

Reading benchmarks and thresholds are numerical measures of proficiency in specific reading skills, that may be used to monitor whether children are on track. Benchmarks and thresholds can inform a shared vision of what successful reading looks like at the end of grade 1, 2 and 3. They provide a standard against which teachers can measure learners' reading subskills and identify *early on* learners who are at risk of not learning to read for meaning by age 10. This, in turn, supports remediation at an earlier age. Additionally, as specific learners reach different benchmarks, this can help teachers adapt their instructional focus to meet the learners' needs at their reading level.

While reading benchmarks exist in English (Hasbrouck and Tindal, 2006), there is scant research guiding the development of Nguni language reading benchmarks or thresholds. One cannot simply transfer reading benchmarks from English to the Nguni languages due to differences with the phonological, morphological and orthographical features of African languages. Benchmarking processes need to take account of the linguistic features of the language for which the benchmarks are being developed. At the simplest level, it makes no sense to compare fluency across languages with vastly different word lengths. Beyond that, one needs to allow for language specific accuracy-speed and fluency-comprehension relationships that reflect reading development.

Method

Our approach to establishing benchmarks builds on a theoretical understanding of reading development. The theoretical framework (explained in section 3) then sets the foundation for an exploratory analysis of available early grade reading assessment (EGRA) data in three Nguni languages. Drawing on learner assessment data from five different studies, we compile the largest available source of information on early grade reading in these languages. When combined, these data contain multiple assessment points which can be compared over time for nearly 16,400 learners in the early grades. This data sample is not nationally representative. However, it provides

a very clear indication of the foundational reading skills of learners in predominately no-fee schools (660 different schools), located across 4 provinces and in relatively high poverty contexts.

A key decision in benchmarking is selecting which reading subskills should be benchmarked. We focus on two: letter-sound knowledge which refers to alphabetic knowledge of the written code; and oral reading fluency (ORF). ORF refers to the ability to read words in context with speed, accuracy and prosody.

♦ Accuracy reflects the percentage of words that are read correctly.
♦ Speed reflects the number of words that are attempted in a time period.
♦ Prosody reflects how natural reading sounds (how it conforms to speech rhythms and intonation patterns and reflects punctuation conventions).

Since assessment of prosody is subjective and it is difficult to measure in field studies, measures of ORF typically focus only on speed and accuracy. In this report, the term *fluency* is used to describe reading with speed and accuracy and is typically measured by the number of words correctly read per minute from a passage of text.

Our approach to benchmarking ORF and letter-sounds is based on a conceptualisation of different stages of reading development where different cognitive processes come into play as reading proficiency increases. Within each process, accuracy develops first followed by speed. Our approach aligns with the decoding threshold hypothesis put forward by Wang et al. (2019) where reading comprehension is unlikely to develop until decoding exceeds a lower bound threshold level. There may also be an upper threshold, beyond which there are no additional gains (in comprehension) for increasing decoding skills. This suggests that the relationship between fluency and comprehension will break down at low and high levels of fluency.

Accordingly, our analysis of EGRA data seeks to establish if there are regular patterns and trends in the speed-accuracy and fluency-comprehension relationships across studies, languages, grades, and reading passages. We further establish the validity of these benchmarks and thresholds by examining learners' future reading proficiency levels if they had met specified thresholds or benchmarks at earlier grade points.

Reading norms

As a first step, reading norms and trends in decoding skills across grades were established using this pooled data sample of 16,400 learners. We highlight the following findings:

Letter-sound knowledge: Far too many learners are entering grade 1 with no letter-sound knowledge, despite most having attended grade R. We also identify that 1 in 10 learners in this sample are still unable to sound one letter-sound correctly by the end of grade 3.

Knowledge of complex consonants: Learners also experience significant difficulty in reading complex consonant sequences (examples include *hl, dl, kh, tsh, ndl, gcw, ntsw*) in grade 1 and 2. These sounds feature regularly in Nguni languages and knowledge of these more complex consonant sequences is necessary to read most grade 1 level texts in these languages. Thus, mastery in reading these is required early on. This may require an adjustment to the Nguni language curriculum, which does not include the teaching of complex consonant sequences in grade 1.

Oral reading fluency: The analysis reveals slow but steady increases over time in oral reading fluency. By the end of grade 3, on average, learners can read 19-25 correct words per minute (cwpm) from a passage. Sadly, between 15% and 26% of learners in these samples are unable to read one word correctly by the end of grade 3.

Our analysis shows that learners' performance in these various reading subskills were closely correlated, particularly in lower grades. But we note the following exception:

Oral reading comprehension: We observe low levels of comprehension even at higher levels of fluency (correct words per minute) from grade 3 upwards. In other words, some learners can read but can't effectively comprehend what they are reading. This suggests there is not enough focus on teaching learners in Foundation Phase the skills needed to answer oral and written comprehension questions.

Reading benchmarks and thresholds

Within each reading subskill we found remarkably consistent patterns between speed (number of words/letter-sounds attempted in a minute) and accuracy of reading (% of words /letter-sounds read correctly out of those attempted). Accuracy initially increases rapidly as speed increases, but then flattens and levels off once learners achieve about 95% accuracy (i.e. learners read 95 of every 100 words/letter-sounds attempted correctly). For letter-sounds this levelling off in accuracy occurs at speeds of approximately 40 sounds attempted per minute. For reading words in a connected text, the levelling of accuracy occurs at 22-34 words per minute.

All of this analysis, together with expert opinion, allows us to identify the following thresholds and benchmarks:

By the end of grade 1 all learners should be able to read 40 letter-sounds correct per minute.

◆ This appears to be a good early predictor of oral reading fluency (ORF) later in the Foundation Phase.
◆ There are few benefits for improving letter-sound knowledge and speed beyond this point.
◆ Once learners have achieved this level of letter-sound knowledge, decoding instruction should focus on helping learners apply word attack strategies and develop fluency.

By the end of grade 2, all learners should be able to read at least 20 correct words per minute (we have termed this the lower ORF threshold).

◆ This is a *minimum* threshold. If learners do not reach this level of fluency, higher order reading skills are very unlikely to develop. Below this threshold we find little evidence that learners can comprehend what they have read, evidenced by very low oral comprehension or written comprehension scores.
◆ By the end of grade 3, between 53% and 76% of the learners in this sample had reached this grade 2 threshold.

By the end of grade 3, all learners should be able to read 35 correct words per minute.

◆ This acts as an upper fluency threshold.
◆ At this level of fluency reading comprehension becomes increasingly possible when learners read on their own.
◆ Once learners reach this level of fluency, it appears that poor comprehension skills become the limiting factor to further literacy development.
◆ Therefore, at this milestone teachers should focus on teaching learners the skills and strategies needed to tackle written comprehension while encouraging vocabulary and language development.
◆ Fluency skills should continue to improve from this milestone.
◆ Approximately only a quarter of learners in the EGRA studies analysed had reached this fluency benchmark by the end of grade 3.

Figure E 1: Reading benchmarks and thresholds for early grade reading in Nguni languages
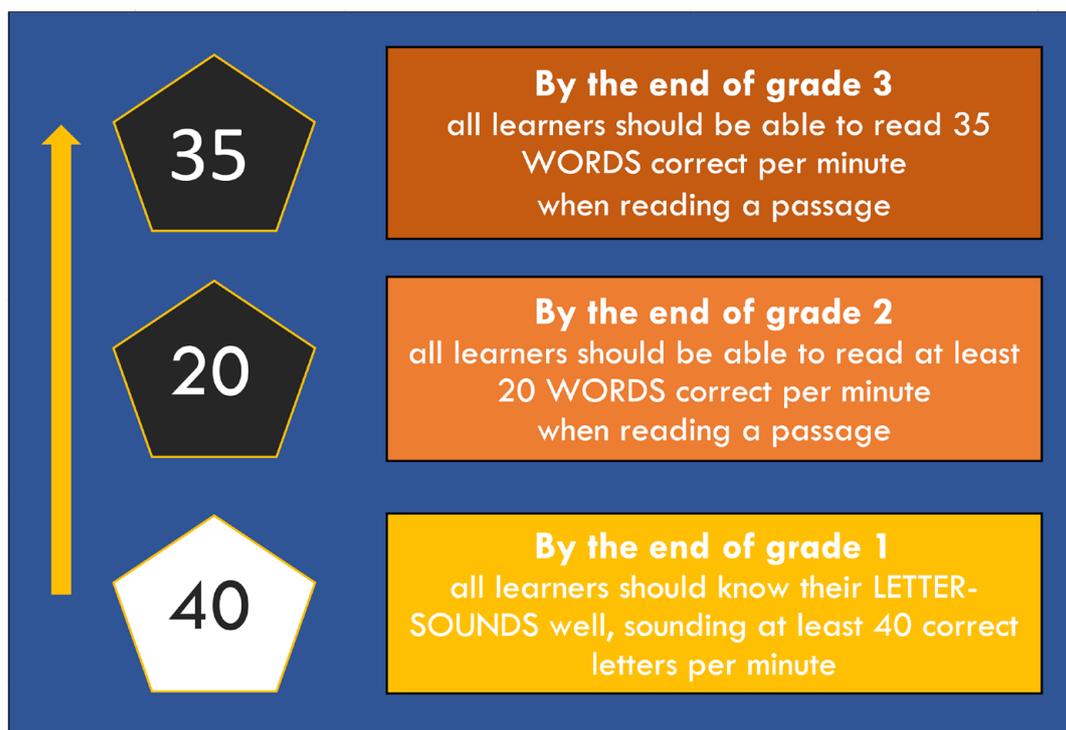
Further research is required

This research is the first of its kind in this context and, as a result, is exploratory. We hope that as more data on early grade reading assessment (EGRA) becomes available, these benchmarks and thresholds will be further tested and corroborated. In particular, we hope that more EGRA data on isiNdebele will be forthcoming to allow the testing of these benchmarks for that language. In the meantime, we suggest that the language structure is similar enough to the three languages tested so that these benchmarks can also be adopted for isiNdebele.

Language benchmarks and thresholds are language- and context-specific. These thresholds apply to the South African context only and only to learners learning to read the Nguni languages.
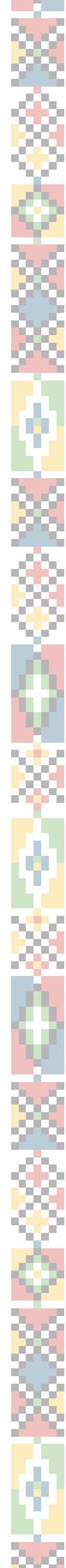
1. Introduction

In early 2019, acknowledging the reading crisis that faces the nation, South African president Cyril Ramaphosa articulated a new and clear expectation for basic education in South Africa: *every child should be able to read for meaning by age 10* (the end of the Foundation Phase, grade 3). Yet in moving towards this goal, various intermediate steps need to be established. This report aims to contribute some milestones that can be tracked in the journey towards the goal of *read for meaning by age 10*, in the form of early grade thresholds and benchmarks for foundational reading skills in Nguni languages.

Reading benchmarks are numerical measures of proficiency in specific reading skills, that may be used to monitor whether children are on track. They are distinct from goals and targets. Goals are often aspirational and communicate a vision of a future outcome, such as: *every child should be able to read for meaning by age 10*. Targets communicate plans to reach these goals within a specific timeframe, such as: *80% of all children should be able to read for meaning by 2030*. In contrast, a benchmark may be: *40 letters correct per minute*. A benchmark such as this would be a way to communicate the speed and accuracy required for mastery of letter-sounds (one of the subskills that contribute to reading for meaning).

Reading thresholds and benchmarks are not decided in an arbitrary manner. They should be informed by scientific literature on reading across language groups. They should be based on strong empirical work and should be sensitive to current realities of learning and curriculum requirements. The use of data to establish benchmarks in this study is guided by language and reading theory, as well as expert advice. We establish benchmarks that are necessary to get learners onto a successful reading trajectory without being too aspirational that no-one can reach them.

This research is an exploratory analysis of early grade reading assessment (EGRA) data in three Nguni languages: isiZulu, isiXhosa and siSwati. This is the largest available dataset on early grade reading in these languages and contains multiple reading assessments for nearly 16,400 learners in early grades. These data allow us to provide the first tentative estimates of early reading thresholds and benchmarks in these three languages.

2. Rationale for developing benchmarks

Reading is a complex phenomenon. Numerous knowledge bases, skills and processes that underpin reading comprehension need to develop and coordinate for learners to become proficient at comprehending text when they read on their own.

The most recent large-scale Progress in International Reading Literacy Study (PIRLS, 2016) showed that 78% of grade 4 learners in South Africa, mostly tested in their home language, could not reach the low international reading benchmark (Howie et al. 2017). While assessments such as PIRLS can identify that learners are not proficient in reading comprehension, they cannot tell us which reading subskills pose problems for learners who struggle to read. They provide little guidance on the subskills of reading that should be targeted in the classroom. The national school curriculum guides the expected competencies learners are required to master in each grade. However, it is silent on how to measure a learner's achievement in specific reading subskills in African languages. To prevent learners from falling behind in the developmental sequence of reading (further described in the next section), we need a shared vision of what reading success looks like at each grade level. The thresholds and benchmark set in this report contribute to this shared vision, as illustrated in Table 1.

Table 1: The uses of reading benchmarks at various levels of the education system

System levels	Clarifying goals and expectations for reading	Clarity & alignment for the effective assessment of reading	Clarifying intervention priorities
NATIONAL AND PROVINCIAL ADMINISTRATION	Benchmarks articulate and communicate an education system's definition of reading proficiency. This promotes alignment of goals across the system with Goal 1 of the DBE Action Plan to 2030.	With a shared understanding of expectations for reading proficiency, the reading assessment process is unified across the system.	With a shared understanding of the size and extent of reading gaps across the system, this paves the way for a constructive intervention response.
SCHOOL	Benchmarks establish standards and targets that school leaders can aim towards and are aligned with national goals for reading proficiency.	Standardises assessment practices across and within schools and ensures that school level assessment is aligned with national goals for reading proficiency.	Clarifies the extent of remedial support required in specific schools and required budget allocations. Focuses intervention responses on the improved teaching of reading and the provision of reading support materials.
CLASSROOM	Benchmarks establish standards and targets that teachers and students can aim towards.	Teachers can determine how many children in their class are on track with their reading.	Target remedial programmes at learners at risk of not being able to read.
HOUSEHOLDS	Benchmarks establish standards for parents against which to assess their children's reading proficiency.	When assessment is linked to standards, and communicated clearly in school reports, this provides meaningful information to parents on how well children read. They can engage in their child's journey to reading proficiency.	Parents and communities are empowered to identify if schools are providing necessary opportunities for their children to learn to read and to partner in remedial programmes.

3. Theoretical framework

While reading for meaning is the goal of reading, a myriad of foundational skills is required for reading and understanding a text on one's own. These skills become integrated through a developmental sequence and it is only gradually that learners come to fully comprehend what they are reading. As illustrated in Figure 1, different processes come into play at different stages of development. The importance of some earlier processes diminish as proficiency increases and they are replaced by qualitatively different processes. Across all the components, accuracy develops first, followed by increased processing speed, which then leads to automaticity (processing without effort or conscious attention)[9]. Reading with fluency frees up cognitive resources (working memory, attention) to be allocated to higher level demands needed for reading comprehension. But the relationship between fluency and reading comprehension does not only go in one direction. Children who understand what they read tend to read more fluently.[10]

Figure 1: Developmental cline in early reading

Phonemic awareness & letter-sounds	Syllable reading	Word reading	Context Fluency (ORF)	Comprehension (literal/inferential/integrative) & metacognition
accuracy	increased processing speed	automaticity	working memory freed for meaning	

Box 1: What is oral reading fluency?

Oral reading fluency (ORF) refers to the ability to read words in context with speed, accuracy and prosody.

♦ Accuracy reflects the number or percentage of words that are read correctly.
♦ Speed reflects the number of words that are attempted in a time period.
♦ Prosody reflects how natural reading sounds (how it conforms to speech rhythms and intonation patterns and reflects punctuation conventions).

As assessment of prosody is subjective, it is difficult to measure in field studies so measures of ORF typically focus only on speed and accuracy. In this report, the term fluency is used to describe reading with speed and accuracy and is typically measured by the number of words correctly read per minute from a passage of text.

Oral reading fluency (see Box 1) has been described as the bridge between decoding and reading comprehension (Pikulski & Chard 2006). Theories of reading arrive at different views on the nature of the intricate relationship between decoding and comprehension.

---

[9] Fuchs et al. 2001; Spear-Swerling, 2006.

[10] The theoretical framework guiding this research is outlined in detail in Section 3 of the technical report. We also note that the development cline does not mean that reading instruction should follow this order. CAPS recommends a balanced approach where decoding and comprehension skills are developed in tandem. Activities such as Listening and Speaking, Shared Reading, Group Guided Reading, Paired Reading and Read Alouds help children develop language, comprehension and thinking skills while Phonics is assigned to the development of decoding skills, supported by Handwriting activities. Group Guided Reading can also be used to develop decoding and fluency skills in the early grades. As mastery in decoding increases, the need for targeted phonics instruction diminishes.

3.1. The Simple View of Reading

The basic premise of the simple view of reading (SVR) is that reading comprehension can be predicted by the following formula: Reading Comprehension = Decoding × Language Comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990; Tunmer & Chapman, 2012). The multiplicative nature of this formula means that if children lack decoding skills (the ability to read words) they would struggle with reading comprehension. Similarly, if they are poor in language comprehension skills (such as oral vocabulary) they would also struggle with reading comprehension. In the early stages of reading, decoding is predicted to have more influence, since children need to master the written code. As decoding skills (for example, word reading or fluency) approach mastery level, their effects on reading comprehension generally decrease, and language comprehension plays a larger role in determining reading comprehension (Fuchs et al, 2012; Caravolas et al. 2013; Garcia & Cain 2014; Torppa et al. 2016).

3.2. The decoding threshold hypothesis

While the SVR assumes a linear relationship between decoding and reading comprehension, the decoding threshold hypothesis of Wang, Sabatini, O'Reilly, and Weeks (2019) argues that the relationship is more complex. They posit 'threshold conditions' in the relationship between decoding and reading comprehension. What this means is that the relationship between decoding and reading comprehension is only observed when decoding occurs above a lower bound threshold level. For learners who decode below the threshold, reading comprehension is unlikely to develop unless decoding can be improved to a level above the decoding threshold (Wang et al. 2019). The authors speculate that there may also be an upper threshold, beyond which there are no additional gains for increased decoding skills. For example, extremely fast decoding does not improve reading comprehension and may cause gaps in understanding.

3.3. Developmental continuum of reading

According to the sequencing of skills, described above, different types of readers can be distinguished on a continuum.

Non-readers show poor print awareness, poor phonological awareness[11] and have very low letter-sound knowledge (and alphabetic awareness) and immature handwriting. Their ability to read words correctly (accuracy) is minimal. Performing below a minimal letter-sound threshold impedes their ability to decode text.

Emerging readers have developed phonological awareness and acquired some basic knowledge of letter-sounds to enable them to blend letters to form syllables or words. Accuracy is increasing, and with it, increased processing speed to read words in or out of context. However, reading is still halting and effortful, and chunking of words into meaningful phrases is not yet regular. Reading comprehension is limited when they read a text on their own.

Readers in the next phase are **developing proficiency.** They have more accurate and fluent knowledge of the alphabetic code, which enables them to decode syllables and words in or out of context with greater accuracy (approximating 95% accuracy which refers to correctly reading 95 of 100 words attempted from a passage). Their processing rate increases to a point where some words are read automatically, and they move beyond the level of sounding out words to articulating meaningful phrases. While their decoding skills are not yet fully automatised, they have freed up enough working memory to construct basic meaning from what they read to support comprehension.

---

11 Phonological awareness is a broad skill that includes identifying and manipulating units of oral language – parts such as words, syllables, and onsets and rhymes. Phonemic awareness refers to the specific ability to focus on and manipulate individual sounds (phonemes) in spoken words.

Competent readers have reached a stage where decoding is accurate (at least 95% accuracy) and largely effortless. Their reading rate is quite advanced for their grade level and they read sentences with natural intonation or prosody. They can read texts containing more complex language and less familiar words; they engage more actively with the text and understand much of what they read. They can respond to questions requiring both integrating information from a specific place in the text (local) with a wider (global) view of the text. Reading becomes a tool for learning – they start learning new things when reading on their own, without mediation from a teacher/adult. They will reread a section of text if comprehension breaks down.

Skilled readers read words in and out of context accurately, effortlessly and quickly, seldom making decoding mistakes. Their reading is automatised; they chunk words into meaningful phrases and construct and integrate meaning. They are equally good at making local and global inferences across the text. The ability to 'read to learn' comes naturally and they will often voluntarily read for information or pleasure. They readily pick up inconsistencies in a text or discrepancies in perspective.

Figure 2 loosely maps the reading skill cline against grade progression in the early school years.

Figure 2: Developmental continuum in early reading in relation to formal grades

FOUNDATION PHASE				INTERMEDIATE PHASE	
Grade R	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**	**Grade 5**

Non-reader	Emerging reader	Developing proficiency reader	Competent reader	Skilled reader
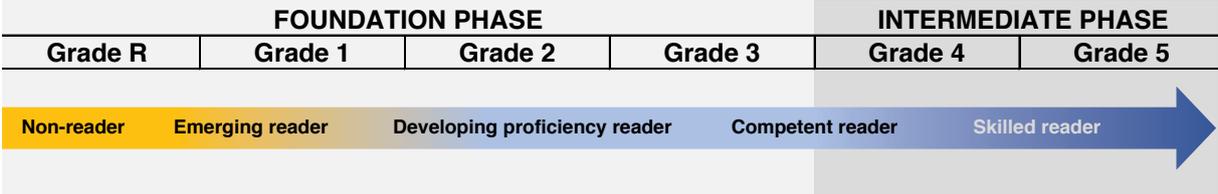
By the time learners exit the Foundation Phase, most of them should be competent readers or at least transitioning from *developing proficiency* to *competent reading*. This developmental sequence shows a general trend; but there will always be exceptions. For example, there may be *emerging readers* in grade R and *competent readers* in grade 2, or a few *non-readers* at the end of grade 1. However, there should not be any *emerging readers* at the end of grade 3. Such a developmental lag would point to challenges in pedagogy and the teaching/learning context.

4. THE NEED FOR LANGUAGE SPECIFIC BENCHMARKS

Benchmarks for reading exist in various other languages and countries. Particularly well-defined ORF benchmarks have been established for different grades for reading in English (Hasbrouck and Tindal, 2006). But, until now, detailed work on the development of reading benchmarks in the African languages has not been done in South Africa. One cannot simply transfer reading benchmarks from English due to differences in the phonological, morphological and orthographical features of African languages. These differences are highlighted in Table 2.

Early reading development differs from one language to the next. In transparent codes (such as Finnish or Nguni languages), accuracy is reached relatively early. This may result in language comprehension being more strongly related to reading comprehension than decoding at earlier grades than for opaque languages (such as English). The straight-forward nature of letter-sound knowledge in transparent languages may allow learners to achieve mastery of this foundational skill earlier in their reading journey.

However, other features of Nguni languages may delay mastery of basic decoding skills. Nguni languages are agglutinating languages where prefixes or suffixes are "glued" to a word stem to form new meanings. Nguni languages also have a conjunctive orthography with morphemes (the smallest meaningful unit of language) merged together into single written words. This results in both long words and a high degree of visual similarity within and between words, making decoding more challenging. Not all African languages do this. For example, the Sotho languages (like Setswana or Sepedi) represent these meaningful units orthographically as separate, smaller words rather than very long words. They have a 'disjunctive' orthography. Like other African languages, Nguni languages also include a large number of complex consonant sequences (examples include *hl, dl, kh, tsh, ndl, gcw, ntsw*) and they have a larger code set than English (isiZulu and isiXhosa have about 60 phonemes compared to 44 phonemes in English).

Benchmarking processes clearly need to take account of the linguistic features of the language for which the benchmarks are being developed. At the simplest level, it makes no sense to compare fluency across languages with vastly different word lengths (see Spaull, Pretorius & Mohohlwane, 2020). Beyond that, one needs to allow for language specific accuracy-speed and fluency-comprehension relationships that reflect reading development.

4.1. Which languages are these benchmarks for?

The aim of this project is to analyse existing large EGRA data sets on early language and reading in three African languages in the Nguni language family (isiZulu, siSwati and isiXhosa) to provide a clearer view of how reading development unfolds in these languages. Then we use these findings to inform decisions on benchmarking. isiNdebele is also a Nguni language, however no data was available, so isiNdebele could not be included in this work.

Table 2: Some differences between Nguni languages, Sotho languages and English

NGUNI LANGUAGES	SOTHO LANGUAGES	ENGLISH
Agglutinating Affixes and word stems are "glued" together to form new meanings. E.g. isiZulu: Bakufundisile Stem: fund	**Agglutinating** Affixes and word stems are "glued" together to form new meanings. E.g. Southern Sotho: ba le rutile Stem: rut	**Analytic** Use of helper words, prepositions, and word order to convey meaning. E.g. They taught you.
Transparent/ Shallow There is a one-to-one letter to sound correspondence. E.g. 'g' is pronounced the same in "ngoko" and "jonga" in isiXhosa	**Transparent/ Shallow** There is a one-to-one letter to sound correspondence. E.g. 'g' is pronounced the same in "gape" and "morago" in Sepedi	**Opaque / Deep** The same letter can represent different sounds in different words. E.g. the sound 'g' differs in the words: 'gate' and 'germ'
Conjunctive Morphemes (the smallest meaningful unit) are merged together into single written words. E.g. isiZulu: Ngiyabathanda	**Disjunctive** Morphemes mostly appear as single words. E.g. Sepedi: ke a ba rata	**Disjunctive** Morphemes mostly appear as single words. E.g. I like them

5. Methodology

Our approach combines rigorous analysis of learner performance data with expert understanding of learners' reading development and the curriculum requirements. In this section we outline the data used and the empirical approach adopted in establishing benchmarks. The benchmarking approach draws on theoretical understandings of the reading development process as described earlier.

5.1. Data

We combined EGRA data from five different reading and literacy studies in South Africa: the Story Powered Schools (SPS) evaluation; the Early Grade Reading Study II (EGRS II); the Funda Wande (FW) evaluation; the Leadership for Literacy (LFL) project and the Zenex Foundation Literacy Project (Zenlit). These data were collected between 2016 and 2019 and assessed learners at some point between grades 1 and 6[12]. All five studies collected EGRA in isiXhosa, isiZulu, or siSwati (Ardington and Meiring 2020; Ardington, Hoadley & Menendez, 2019; Department of Basic Education 2019; Taylor, Wills & Hoadley, 2019; Zenex Foundation, 2018).

5.1.1.   Reading measures across datasets

The dataset was used in its pooled form in a preliminary analysis to identify readings norms and trends across grades for letter-sound knowledge, oral reading fluency and reading comprehension. Sub-sets of the pooled data are then used in the benchmarking analysis. This analysis focuses on examining interactions between these three reading subskills.

Assessment data were collected on letter-sound knowledge for grades 1 to 4 learners, single word reading for grades 1 to 5 learners, and oral reading fluency (ORF) and ORF comprehension for learners from grade 1 to 6. In one study (FW) complex consonant sequences are assessed separately as distinct tasks. All studies (except Zenlit) include item level data. All were conducted as one-on-one assessments. For more detail of the EGRA subtasks tested by grade, term and language, please refer to the technical report.

Sub-sets of the pooled data, including reading comprehension performance data are then used in the benchmarking analysis.

5.2. Data samples used for establishing reading norms

When pooled, this is the largest, most up-to-date source of early reading performance data in Nguni languages that exists, to our knowledge. Almost 16,400 unique learners (with up to four assessments at different grade-term points) in more than 660 schools in four provinces (Eastern Cape, KwaZulu-Natal, Gauteng and Mpumulanga) are represented in the data. Sample sizes of individual studies vary considerably (as shown in Table 3). The isiZulu and isiXhosa samples are dominated by the Story Powered Schools (SPS) evaluation. The EGRS II study is the only available siSwati sample.

None of these samples were intended to be representative of early grade learners in any province or language group. However, we summarise the types of schools and learners that are most represented:

---

12 The SPS, FW and EGRS II studies were designed as randomised control trials to assess the effectiveness of different literacy programmes on learner literacy in home language or first additional English. The LFL project was a mixed-methods study to assess the functioning of township and rural schools and to explore how leadership contributes to literacy learning environments. Zenlit aimed to establish the efficacy of literacy programmes to support the teaching of reading.

- ◆ Almost all learner reading outcomes (99%) are obtained from learners in non-fee charging (also called quintile 1 to 3) schools[13];
- ◆ 86% of learners across the pooled data are in rural schools;
- ◆ Almost all learners in the sample (over 98%) were tested in a Nguni language which matched the Foundation Phase language of learning and teaching (LOLT) in their school and their home language;
- ◆ Learner samples tested in isiZulu are in KwaZulu-Natal, Mpumulanga and Gauteng;
- ◆ Learners tested in isiXhosa are all from the Eastern Cape; and
- ◆ Learners tested in siSwati are from one study in Mpumalanga.

Except for the Zenlit study, all studies considered are longitudinal (have more than one assessment point per learner) and track the same learners over time. In all longitudinal data, attrition was below 20%[14]. Data was combined across treatment and control schools to allow enough variation in the reading scores for meaningful analysis.

Table 3: Grades at which EGRA was conducted, by individual study

Study	Grade	Language	No. schools	No. of learners*
SPS	Grades 2-5	isiZulu isiXhosa	188 170	5371 5002
EGRS II	Grade 1-3	isiZulu siSwati	49 131	969 2358
FW	Grade 1 and 2	isiXhosa	59	1187
LFL	Grade 3 and 6	isiZulu	42	510
Zenlit	Grades 1-3 Grades 1-3	isiZulu isiXhosa	12 10	538 450

Notes: *Sample size here reflects the entire sample combined for each study by language and where there are multiple waves of data collection, only wave 1 sample sizes are counted.

5.3. Data sub-samples used for benchmarking analyses

For establishing reading thresholds and benchmarks, sub-sets of the data were used[15]. Table 4 summarises the samples used in establishing oral reading fluency (ORF) thresholds, and ORF and letter-sound benchmarks. All the samples are large and drawn from substantial numbers of schools as shown in Table 3 and 4.[16] Large proportions of the learners can read at least one word in a connected text ranging from 45% at the end of grade 1 to over 80% by the end of grade 3. There is therefore enough variation to detect patterns in accuracy, speed, fluency and comprehension.

Three samples (shown in Table 4, in the column headed "3-minute oral reading") are used in the analysis of oral reading comprehension for benchmarking. Typically, ORF assessments are limited to 60 seconds. The associated comprehension questions are then asked, but only up to the point in a passage that the learner reached in 60 seconds. This limits the ability to assess the association between fluency and comprehension, as slow readers are not asked more than a few questions (Ardington and Menendez, 2020). In three of the study samples learners were timed to 60 seconds and then allowed to read for a further 2-minutes.

---

13 The South African public-school system categorises schools into 'Quintiles' depending on the wealth of the area in which the school is located. The poorest three quintiles of schools do not charge fees, while quintiles 4 and 5 schools are typically fee-charging. Quintile 5 schools tend to serve the more affluent students and are of much higher quality and functioning than schools classified in poorer quintiles.
14 Suggesting sample characteristics do not change significantly at higher grades due to loss to follow-up.
15 The specific sub-set used for each threshold depended on the EGRA subtasks assessed at specific grades, the availability of item-level data and the specific nature of the tasks administered.
16 The control and intervention samples of the individual studies were pooled to increase the sample size and maximize the possible distribution of reading scores.

Table 4: Sample sizes by language, study and grade: Nguni language benchmarking analysis

Language	Grade Term	Study	Passage	Letter-sounds (item level)	ORF (item level)	ORF >0	ORF > 0 (%)	3-minute oral reading	Attempting all comprehension questions	Attempting at least 5 comprehension questions	Written comprehension	Oral reading	Written
												Number of comprehension questions	
isiXhosa	G1T1	FW	1	595	-	-		-	-	-	-	-	-
	G1T4	FW	1	555	555	251	45%	228	104	193	-	12	-
	G2T1	SPS	1	947	946	535	57%	-	-	-	-	-	-
	G2T1	FW	1	592	592	332	56%	322	154	268	-	12	-
	G2T4	FW	1	561	561	468	83%	444	340	424	-	12	-
	G2T4	FW	2	-	561	431	77%	431	263	414	-	10	-
	G3T1	SPS	1	940	939	632	67%	-	-	-	-	-	-
	G3T3	SPS	1	1410	1410	1204	85%	662	466	625	-	10	-
	G4T1	SPS	1	942	942	780	83%	-	-	-	-	-	-
	G4T3	SPS	1	-	1422	1232	87%	689	562	672	-	9	-
	G5T3	SPS	1	-	1452	1315	91%	720	584	654	1428	9	8
isiZulu	G2T1	SPS	1	925	925	598	65%	-	-	-	-	-	-
	G2T4	EGRS II	1	766	765	433	57%	-	-	-	-	-	-
	G3T1	SPS	1	925	925	717	78%	-	-	-	-	-	-
	G3T1	LFL	1	509	509	405	80%	-	-	-	-	-	-
	G3T3	SPS	1	1477	1475	1293	88%	652	510	601	-	10	-
	G3T4	EGRS II	1	762	762	593	78%	593	514	514	757	5	6
	G3T4	LFL	1	430	430	372	87%	-	-	-	-	-	-
	G4T1	SPS	1	925	925	820	89%	-	-	-	-	-	-
	G4T3	SPS	1	-	1534	1404	92%	730	622	711	-	8	-
	G5T3	SPS	1	-	1474	1374	93%	708	560	665	1441	9	8
Siswati	G2T4	EGRS II	1	1995	1994	1252	63%	-	-	-	-	-	-
	G3T4	EGRS II	1	1921	1922	1430	74%	1430	1340	1340	1875	5	6

Where: G1T1 = Grade 1, Term 1; G2T2 = Grade 2, Term 2, etc

5.4. ORF text and ORF comprehension analysis

The texts used to assess oral reading fluency and oral reading comprehension were closely analysed to determine their comparability. The studies drew on just 3 narrative texts. But the details of the texts (and level of difficulty) varied, even when based on the same story. The texts were remarkably similar in terms of how many syllables were reached by the 10th and 30th words. They were also very similar with respect to the number of complex consonant sequences up to the 10th and 30th word. These similarities justify using the various EGRA data together for ORF benchmarking purposes. The technical report provides more detail on the text analysis.

The oral comprehension questions associated with the ORF texts range from easy, literal questions to more challenging inferential, integrative and evaluative questions. Despite variation in the types of comprehension questions asked, they were typically based on narrative texts and tested predominately lower levels of reading comprehension (literal comprehension and straightforward inferences). For more detail on the categorisation of comprehension questions by comprehension process, please see the technical report.

5.5. Our approach to benchmarking

Our approach to benchmarking is based on an understanding of different stages of reading development (outlined in the theoretical framework) where different processes come into play as reading proficiency increases (see Figure 1). Within each process, accuracy develops first followed by speed. Therefore, we explicitly analyse both accuracy and speed and their interrelationship. Our analytical approach aligns with the decoding threshold hypothesis put forward by Wang et al. (2019) where reading comprehension is unlikely to develop until decoding exceeds a lower bound threshold level. There may also be an upper threshold, beyond which there are no additional gains in comprehension for increasing decoding skills. This suggests that the relationship between fluency and comprehension will break down at low and high levels of fluency.

We aim to identify:

♦   A letter-sound benchmark
♦   An oral reading fluency (ORF) threshold
♦   An ORF benchmark

The ORF threshold represents the *minimum* level that must be achieved, to begin engaging higher order skills in the written mode. The letter-sound and ORF *benchmark* represent upper thresholds, above which there are diminishing returns to a continued focus on this skill. The benchmarks should signal to teachers the need to shift the instructional focus to higher order reading skills. For example, when the ORF benchmark has been reached, this signals to teachers to focus on developing comprehension skills and strategies, to support fluency and prosody and to diminish attention to phonics.

Our method relies on visualisation and non-parametric statistical techniques to examine the relationships between accuracy, fluency, and comprehension. We seek to establish if there are regular patterns and trends in the relationship between speed and accuracy, and fluency and comprehension across studies, languages, grades, and reading passages. If regular patterns exist, they allow us to identify critical thresholds, which are independent of the text used and the difficulty of comprehension questions asked.

Once potential thresholds are identified, we test them to determine if they:

♦   Are attainable by current learners, but ambitious enough to lead to meaningful improvement in reading;

- ♦ Correspond to meaningful and distinguishable zones along the reading cline, and
- ♦ Align with the stages of reading development.

Where data were available to follow learners' reading performance over time, the proposed thresholds were examined relative to learners' previous and future reading proficiency scores. This analysis was conducted to determine whether achieving the proposed threshold predicts reading performance in later grades.

The establishment of a letter-sound benchmark relied on expert opinion of letter-sound knowledge required for mastery in decoding and insights from the data. Where data was available at two timepoints, it was used:

i.   to examine improvements in letter-sound knowledge against the learners' baseline score to identify the point beyond which gains in letter-sounds were negligible; and
ii. to investigate whether meeting the letter-sound benchmark is related to future reading proficiency.

The approach to benchmarking adopted here has several advantages over previously used methodologies. First, non-parametric methods make no assumptions about the speed-accuracy or fluency-comprehension relationships. This is necessary as these relationships can be affected by the linguistic differences between languages and pedagogy (in particular, the attention given to teaching comprehension skills[17]). Our analysis follows the developmental cline of reading and communicates to teachers the points at which they should change the focus of their instruction. The key disadvantage of this method is that it requires fairly advanced data analytical skills together with expert judgement and may be difficult to replicate.

> **Box 2: An explanation of why we don't benchmark ORF to a fixed comprehension level.**
>
> Typical benchmarking approaches rely on identifying a comprehension threshold (e.g. at least 80% of questions correct) and then using statistical techniques to identify the fluency levels which are associated with meeting that threshold. These methods are very dependent on the level of difficulty of the comprehension questions and their placement in the text (how far a learner needs to read in the text to attempt these questions). There can be a wide range of difficulty within a single type of comprehension question (eg: literal comprehension questions) making it difficult to assess how comparable passages and sets of questions are. This raises significant methodological challenges as it can make the choice of a cut off for a desired level of comprehension (eg: 80% of questions correct) somewhat arbitrary. For these reasons, we have not adopted this approach to benchmarking.

---

17 A pedagogy that focuses little on teaching comprehension skills can result in lower comprehension scores in that context which are independent of learners' reading skills.

6. Preliminary data analysis: Reading norms and trends

Before presenting the benchmarking process and analyses, we present summary descriptive statistics for key decoding skills across grades 1 to 6 for three languages. We reiterate that the sample is not nationally representative but reflects reading performance for learners in no-fee schools in higher poverty contexts.

6.1. Letter-sound knowledge

There is considerable variation in letter-sound knowledge by grade, term, and across studies (see the technical report for more details). As a result, it is not possible to provide a 'norm' across these data. Nevertheless, some patterns or trends emerge as seen in Figure 3:

♦ Far too many learners are entering grade 1 with virtually no letter-sound knowledge, despite having been in grade R.

♦ Across the data sets, the largest gains in letter-sound knowledge across all samples are acquired in grades 1 and 2. Gains in letter-sound knowledge are muted in grade 3, even for learners coming off a low base. This suggests it is vital that children master this skill in the first two years of school. The opportunities to 'catch-up' this skill beyond the first two grades are limited.

♦ Far too many learners are not mastering this basic decoding skill by the end of grade 3 with around 1 in 10 learners across the pooled data unable to sound one letter correctly at the end of the Foundation Phase.

Figure 3: Letter-sound score by language, grade and term (pooled EGRA dataset)
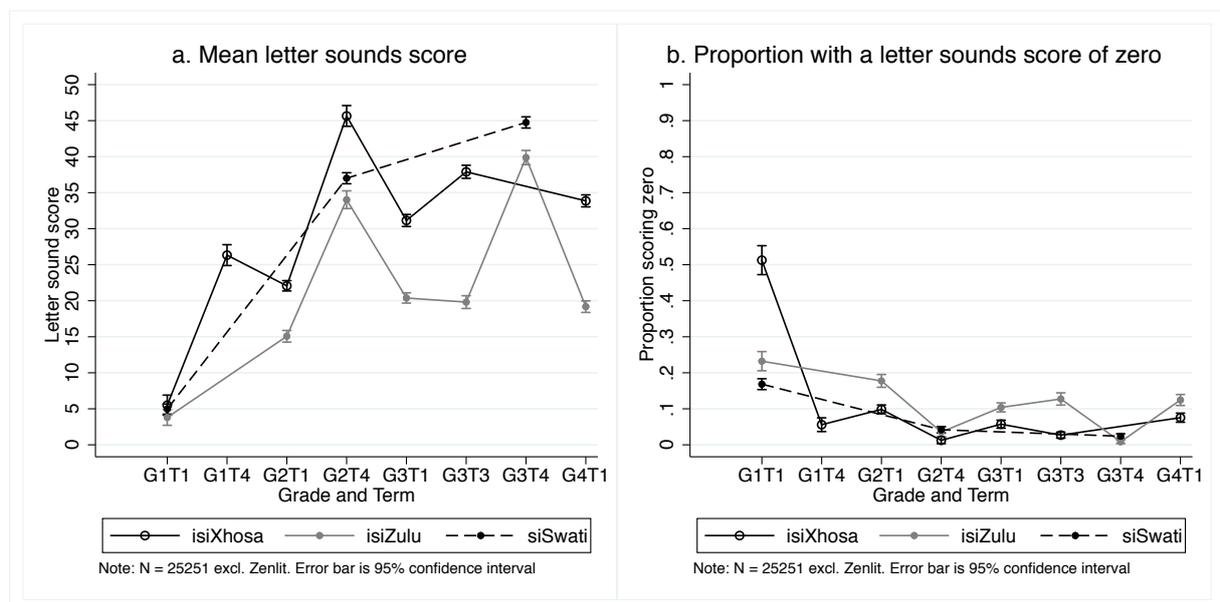
Source: Pooled Nguni EGRA dataset. All letter-sound tasks are timed at one minute except for EGRS II, grade 1. Where G1T1 = Grade 1, Term 1, G1T4 = Grade 1, Term 3 or 4, etc.

6.2. Knowledge of complex consonants

Learners find reading single letter-sounds much easier than complex consonant sequences (examples include *hl, ng, ph, tsh, ndl, gcw*). This is shown in Figure 4 which compares the proportion of learners scoring zero on letter-sounds with those scoring zero on an assessment of complex consonants sequences. By the end of grade 1, virtually all learners in this isiXhosa sample can sound at least one letter, but only half can sound one complex consonant correctly. The lack of familiarity with complex consonant sequences is a major inhibitor to reading. Quite simply, it is not possible to begin reading a passage without mastering these. We show in the technical report that grade 1 texts in Nguni languages require that learners have a knowledge of complex consonant sequences.

Figure 4: Zero scores in single consonant letter-sounds versus complex consonant sequences
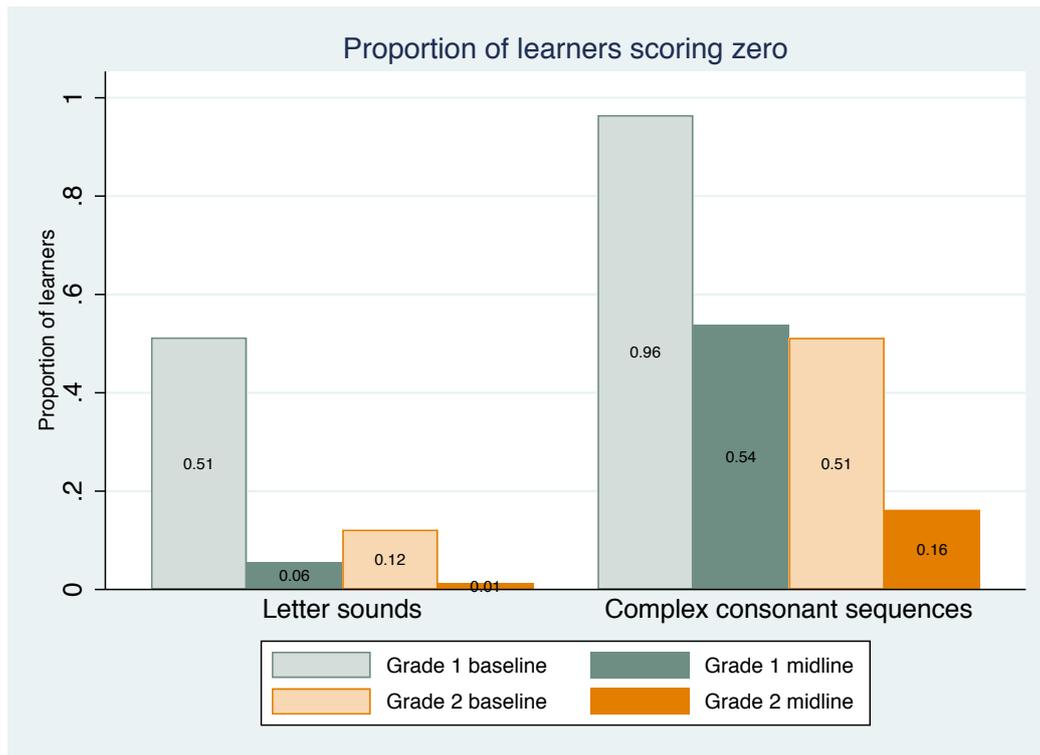
Source: Funda Wande evaluation, isiXhosa sample.

6.3. Text reading (Oral Reading Fluency)

Figure 5(a) shows learners' mean oral reading fluency scores expressed as the number of words correctly read per minute (cwpm) when reading a passage of connected text. These scores are shown for learners in the sample from the end of grade 1 to the end of grade 6. The percentage scoring zero is shown in panel b of the same figure.

♦ Currently, the ability to read text in a passage only really emerges by the end of grade 3.
♦ Around half of the learners in the urban isiXhosa sample cannot read one word correctly in a grade-level text by the end of grade 1.
♦ By the end of grade 2, 20% of the isiXhosa learners, 43% of the isiZulu learners and 37% of the siSwati learners still cannot read one word correctly from a passage.
♦ By the end of grade 3, learners read on average 19-25 cwpm depending on the language in question (19 cwpm for isiXhosa, 22 cwpm for siSwati and 25 cwpm for isiZulu).
♦ By the end of the Foundation Phase 15% of isiXhosa learners, 19% of isiZulu learners and 26% of siSwati learners in these samples are unable to read a single word from a passage, as shown in Figure 5(b).

The norms and trends in learners' mean score for reading familiar words (single words presented out of context (i.e. in a list)) are so similar to those for text reading that, for the sake of simplicity, they are not presented or discussed in this report.

Figure 5: Oral reading fluency score by language, grade and term (pooled EGRA dataset)
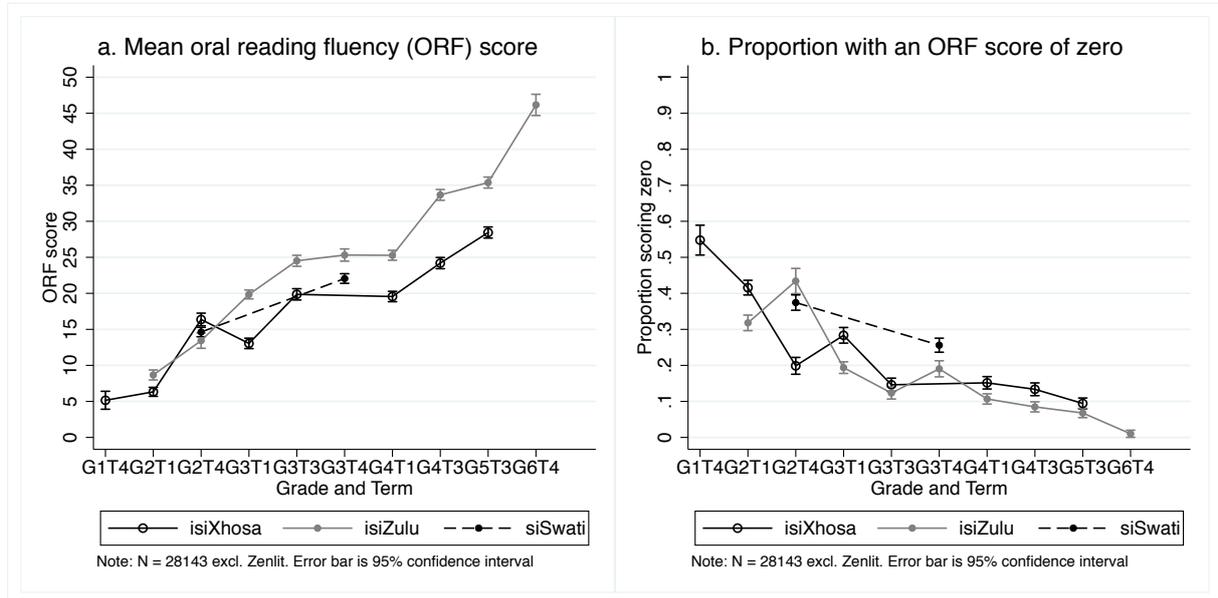
Source: Pooled Nguni EGRA dataset. Where G1T1 = Grade 1, Term 1, G1T4 = Grade 1, Term 3 or 4, etc.

6.4. Comprehension

As shown in the first column of Table 5, from the end of grade 3 at least two thirds of learners in our data samples attempt to read through a whole passage of text presented within 3 minutes. Comprehension scores (for learners who completed reading the passage) range from 48% to 77%. But there is no clear progression in comprehension scores across grades because ORF text passages and comprehension questions differ across samples. As we show in the technical report, there is also no clear relationship between question difficulty and the hierarchy of the comprehension process. We find evidence of considerable variation in difficulty within each comprehension process[18]. This points to the challenges experienced in anchoring comprehension questions to a required level. Measuring proficiency in comprehension is complex.

---

[18] A similar pattern holds in the case of PIRLS. PIRLS classifies question difficulty according to the lowest benchmark (low international, intermediate, high, advanced) at which a specified minimum percent of learners answer questions correctly. For example, a multiple-choice question is classified at the low international benchmark if at least 65% of learners with scores between 390 and 410 answer the question correctly. While most literal questions are anchored at the low international benchmark, there are literal questions at every level of difficulty. Within the other comprehension processes, questions also range in difficulty level.

Table 5: Oral reading comprehension score for learners attempting all comprehension questions after reading for 3-minutes

Language	Grade Term	Learners attempting all questions (%)	Comprehension score for learners attempting all questions (%)
isiXhosa	G1T4	17%	59%
	G2T1	26%	68%
	G2T4	57%	67%
	G2T4	47%	48%
	G3T3	57%	75%
	G4T3	68%	61%
	G5T3	72%	65%
isiZulu	G3T3	66%	77%
	G3T4	67%	68%
	G4T3	78%	69%
	G5T3	73%	69%
siSwati	G3T4	70%	65%

6.5. Relationships between code-based factors

The correlation between learner performance in letter-sound knowledge, single word reading and ORF by grade and term were examined. The correlation between letter-sound knowledge and ORF starts off as very strong but declines at higher grades. The decline suggests that as decoding skills improve, other skills become more important for fluency when reading connected text [19].

Learner performance in single word reading and ORF follow a very close relationship across all grades and languages. As a result, it is unnecessary to set benchmarks for both reading skills. In this analysis, benchmarks and thresholds are not set for single word reading.

---

19 Reduced variation (as mastery increases) at higher grades also limits detectable correlations.
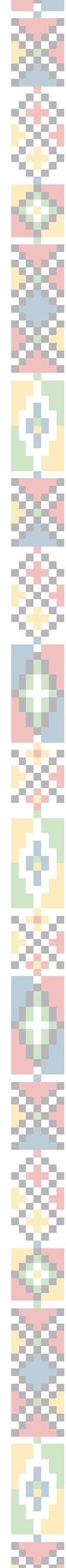
7. Benchmarking

The benchmarking analysis provides more in-depth analysis of ORF, comprehension skills and letter-sound knowledge and the interrelationships between them.

7.1. Establishing fluency thresholds

Our analysis process follows the developmental cline of reading (see Figure 1 in section 3). Thus, we begin with a focus on accuracy, as this needs to develop before increasing speed and automaticity.

7.1.1. Relationship between speed and accuracy

The relationship between speed and comprehension is moderated by accuracy, with errors both reducing speed and cluttering working memory. The accuracy-speed relationship is likely to differ depending on the characteristics of the language; thus we investigate the relationship between oral reading speed and accuracy for each of the Nguni languages in Figures 6 to 8. (The relationship is displayed using polynomial regressions[20]). Reading speed is measured by the number of words attempted (in the time limit) from a passage. Accuracy is measured by the percentage correct of those words attempted. The figures show the average accuracy associated with each level of speed. The figures include grey dashed vertical reference lines at 20 and 35 words per minute and a red horizontal line representing accuracy of 95% (i.e. for every 100 words attempted, the learner gets 95 correct).

Across all languages, grades and reading passages, there is a remarkably consistent pattern: initially accuracy and speed increase steeply together and then the relationship tends to flatten off when accuracy levels reach around 95%. For example, at the beginning of grade 3 isiXhosa learners who are attempting around 10 words per minute are getting every third word incorrect. Accuracy and speed improve steeply and learners reaching 95% accuracy are typically reading at a speed of around 26 words per minute. Thereafter, we see little changes in accuracy with increasing speed, suggesting that an accuracy threshold has been reached. Across all the samples, accuracy of 95% is associated with speeds ranging from 22 to 34 words per minute. The same patterns are observed for all three languages: isiXhosa, isiZulu and siSwati.

Next, we investigated the distribution of speed for accurate readers (learners who could read with at least 95% accuracy) and compared it against the speed of inaccurate learners (who did not achieve 95% accuracy). From grade 3 upwards, very few accurate readers read slower than 20 words a minute. Accurate readers in grade 1 and the beginning of grade 2 tend to be considerably slower. Throughout the Foundation Phase, almost all inaccurate readers read less than 20 words per minute. Almost all inaccurate Intermediate Phase readers read below 35 words per minute[21].

This supports the idea that learners who read less than 20 words per minute have not yet reached accuracy levels to support automaticity. They would benefit from instruction focussed on improving their decoding skills and fluency.

---

20 locally weighted
21 These results are described fully in the technical report.

Figure 6: Reading speed and accuracy, by sample (isiXhosa)

isiXhosa connected text reading: speed and accuracy

Legend:
Grade 1 Term 4, Grade 2 Term 1, Grade 2 Term 1, Grade 2 Term 4, Grade 2 Term 4
Grade 3 Term 1, Grade 3 Term 3, Grade 4 Term 1, Grade 4 Term 3, Grade 5 Term 3
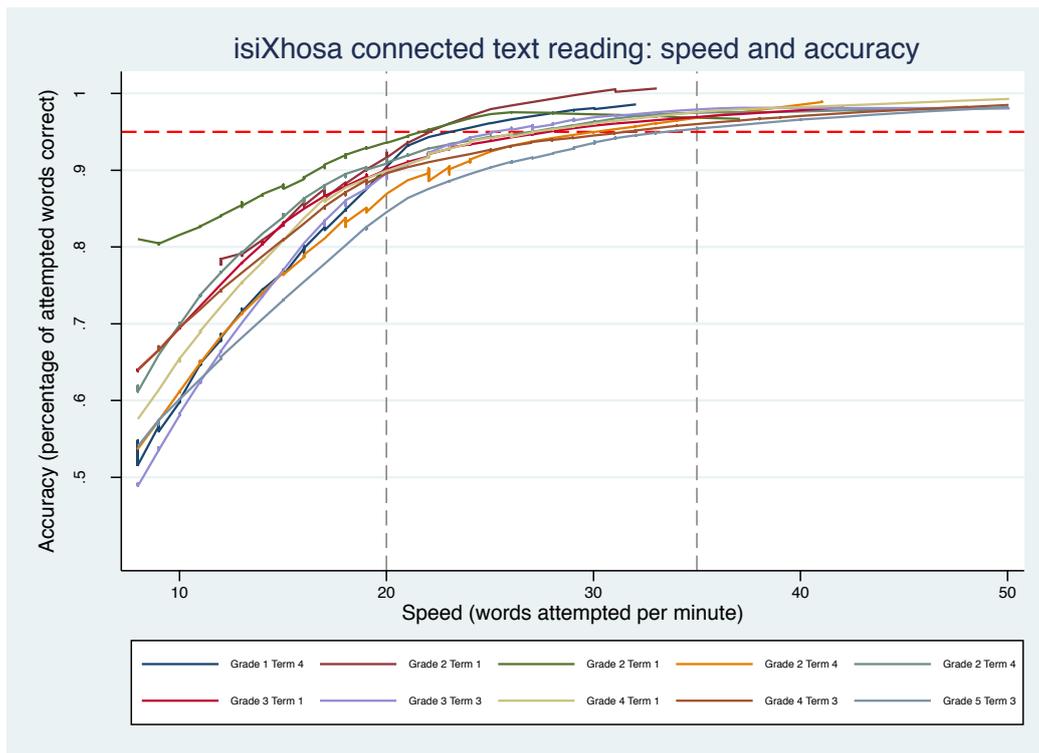
Figure 7: Reading speed and accuracy, by sample (isiZulu)

isiZulu connected text reading: speed and accuracy

Legend:
Grade 2 Term 1, Grade 2 Term 4, Grade 3 Term 1, Grade 3 Term 1, Grade 3 Term 3
Grade 3 Term 4, Grade 3 Term 4, Grade 4 Term 1, Grade 4 Term 3, Grade 5 Term 3
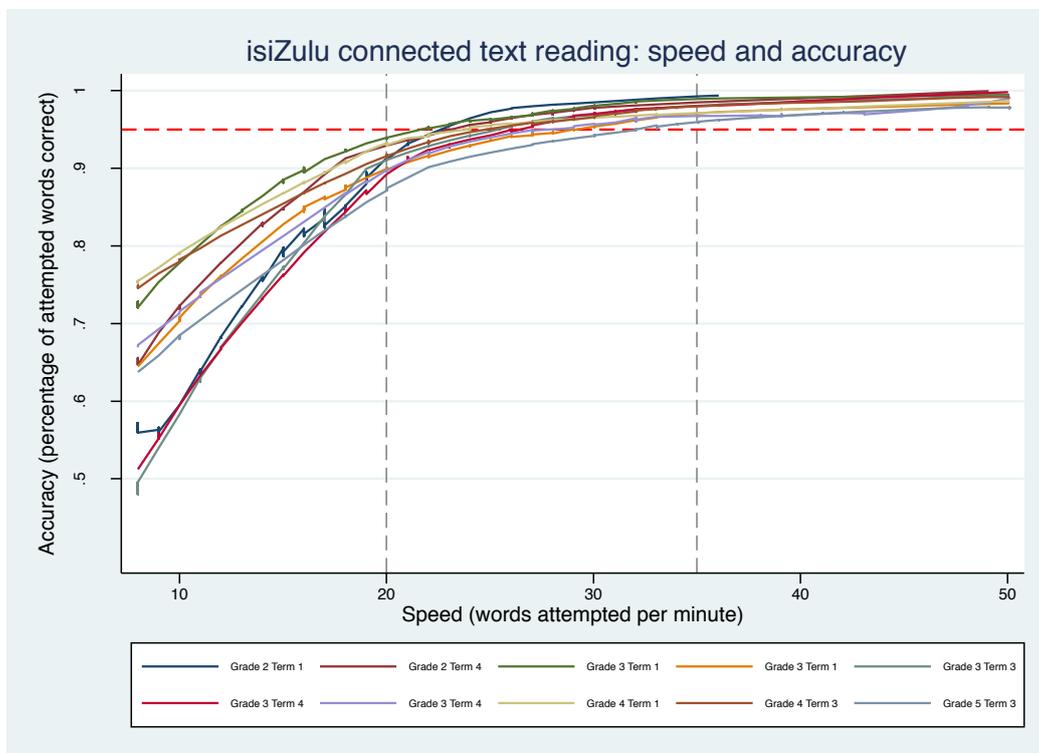
Figure 8: Reading speed and accuracy, by sample (siSwati)
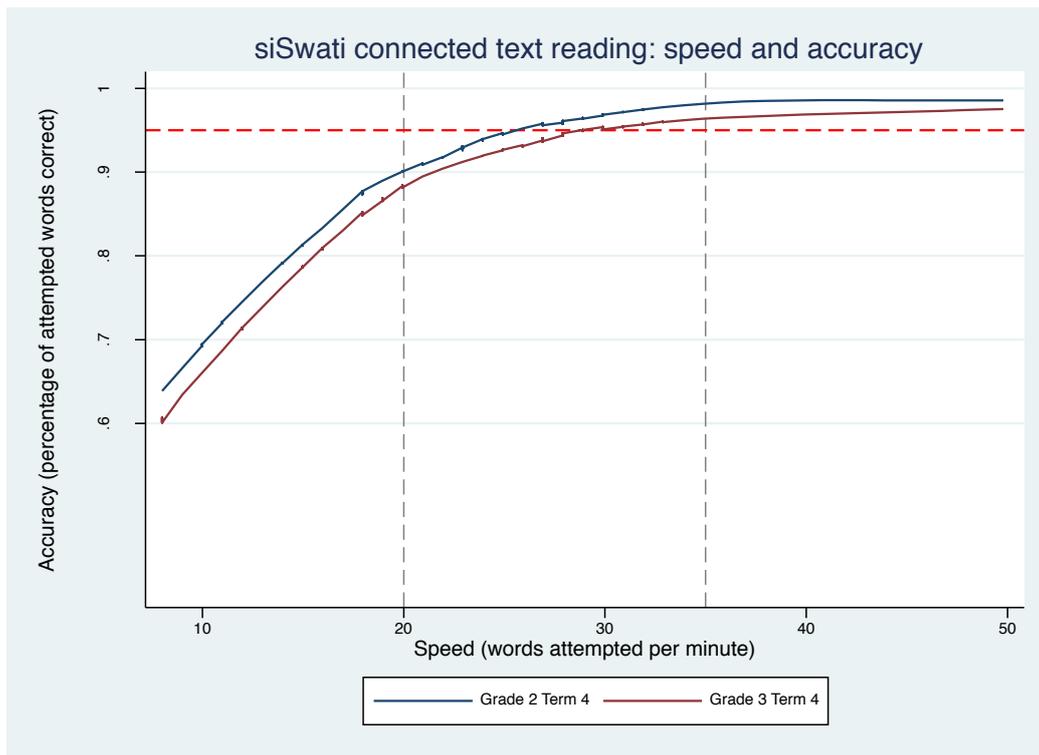
7.1.2.   Relationship between fluency and comprehension

We now examine the relationship between oral reading fluency[22] and oral reading comprehension, among those learners who read far enough within the 3-minute time limit to attempt all the comprehension questions. Figure 9 plots the average comprehension score at each level of fluency separately for all samples that meet these criteria. The figure includes reference lines at 20 and 35 cwpm.

Although there are differences in the average comprehension level between samples, the fluency-comprehension gradient is similar across studies. It is initially very steep and then flattens out. Below 20 cwpm the gradient is very steep. Learners reading below 20 cwpm tend to have very poor comprehension. **_In a separate analysis that compares learners' fluency to their scores on written comprehension tests, those with ORF scores below 20 show little evidence that they can comprehend what they have read (see technical report for more details)._** These learners would benefit from instruction that improves their decoding skills. Fluency of below 20 cwpm appears to be a threshold below which reading comprehension skills are unlikely to develop.

---

22 Fluency, in this context, means the number of correct words per minute from a passage of text.

Figure 9: Relationship between oral reading fluency and comprehension for learners attempting all comprehension questions

Learners attempting all questions

Legend:
- isiXhosa Grade 1 Term 4
- isiXhosa Grade 2 Term 1
- isiXhosa Grade 2 Term 4
- isiXhosa Grade 2 Term 4
- isiXhosa Grade 3 Term 3
- isiXhosa Grade 4 Term 3
- isiXhosa Grade 5 Term 3
- isiZulu Grade 3 Term 3
- isiZulu Grade 3 Term 4
- isiZulu Grade 4 Term 3
- isiZulu Grade 5 Term 3
- siSwati Grade 3 Term 4
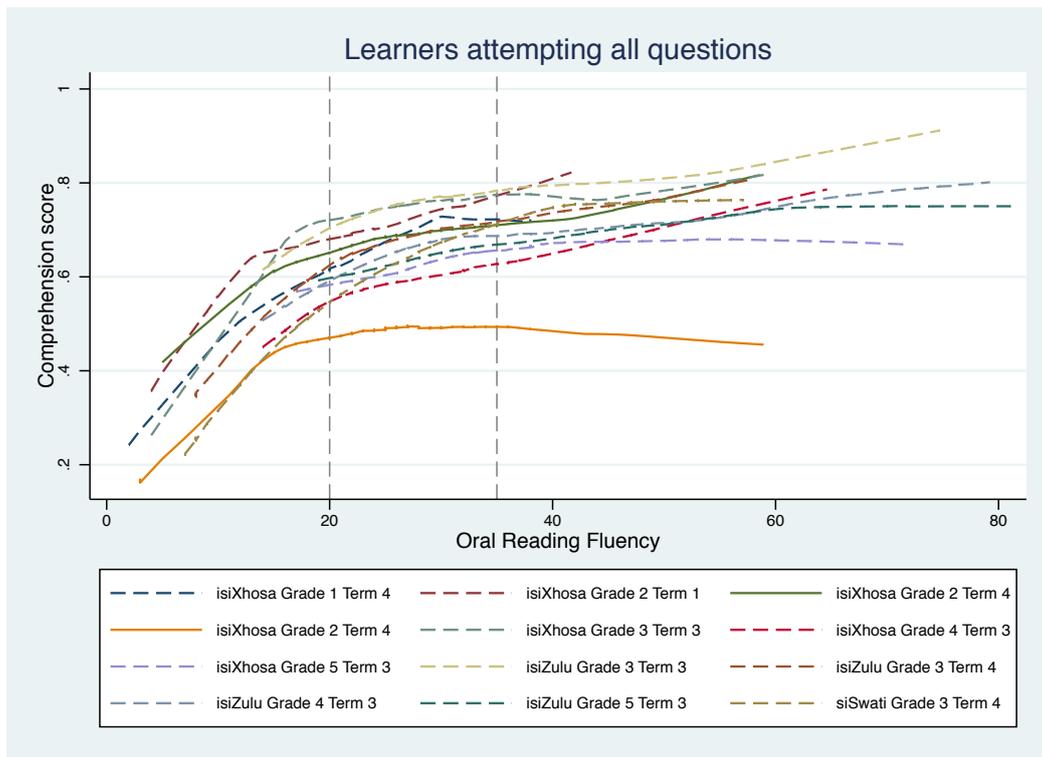
For learners reading between 20 and 35 cwpm, increased fluency is associated with improvements in comprehension. Improvements in fluency above 35 cwpm are associated with smaller improvements in comprehension. The fluency-comprehension gradient flattens at fairly low comprehension levels (between 60% and 80% of comprehension questions correct, in most samples). This supports the idea that there are upper thresholds along the developmental cline of reading. This suggests that underdeveloped comprehension skills become the key hurdle for learners at these levels of fluency. For learners reading more than 35 cwpm, the instructional focus should emphasise teaching comprehension skills.

The variability in the level of comprehension scores between samples highlights the challenge with ensuring equivalence in difficulty across sets of comprehension questions (see the technical report). However, despite these differences in comprehension difficulty, the fluency-comprehension gradient is remarkably consistent. There are clear regular patterns in the data across these Nguni languages, grades and reading passages to support the identification of:

- a lower threshold at around 20 cwpm, below which teaching should focus on improving decoding skills; and
- a higher benchmark at around 35 cwpm, above which teachers' attention should focus on the strengthening of reading comprehension skills.

7.1.3. Fluency thresholds and learner profiles

Figures 10 to 11 categorise learners by language into four fluency groups:

- unable to read one word (shown in blue),
- reading less than 20 words cpm (shown in red),
- reading 20 to 34 words cpm (shown in green) and
- reading at least 35 words cpm (yellow).

Within each language, the rows are organised by grade and term. Although there are differences between studies, the general progression is clear. In grade 1 and 2:

♦ between 23% and 55% of learners are unable to read a single word;
♦ most learners are reading at rates below the lower threshold of 20 cwpm; and
♦ very few learners reach the upper threshold of 35 cwpm.

By the end of grade 3, most learners (53% to 76%, depending on the sample) have reached the lower threshold (20 cwpm) and approximately a quarter have reached the benchmark (35 cwpm)[23]. This suggests that this lower threshold and benchmark (or upper threshold) are not completely out of reach for Foundation Phase learners.

Figure 10: Early grade fluency profiles, isiXhosa samples

isiXhosa oral reading fluency

Grade	Cannot read	Less than 20 CWPM	20-34 CWPM	35+ CWPM
Grade 1 Term 4	55	34	10	1
Grade 2 Term 1	43	50	6	1
Grade 2 Term 1	44	42	13	1
Grade 2 Term 4	17	38	34	12
Grade 2 Term 4	23	38	30	8
Grade 3 Term 1	33	42	22	3
Grade 3 Term 3	15	33	37	16
Grade 4 Term 1	17	32	41	10
Grade 4 Term 3	13	23	37	26
Grade 5 Term 3	9	15	42	33

Percentage of learners

---
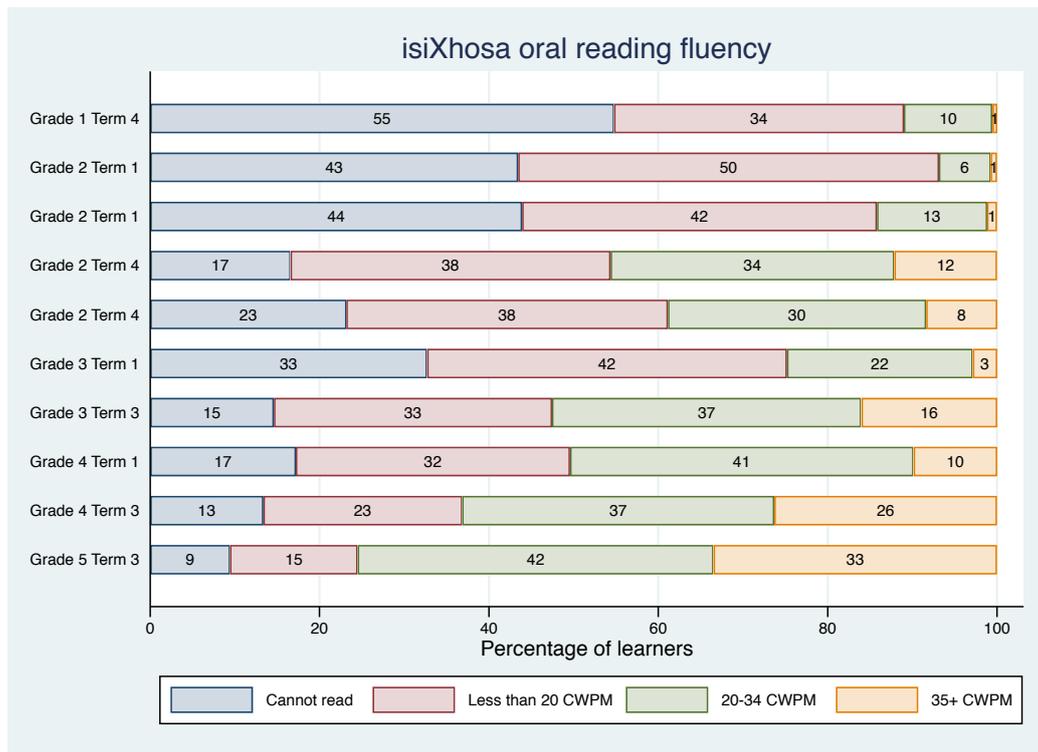
23 This varies from 16% to 47% depending in individual samples.

Figure 11: Early grade fluency profiles, isiZulu and siSwati samples

isiZulu and siSwati oral reading fluency
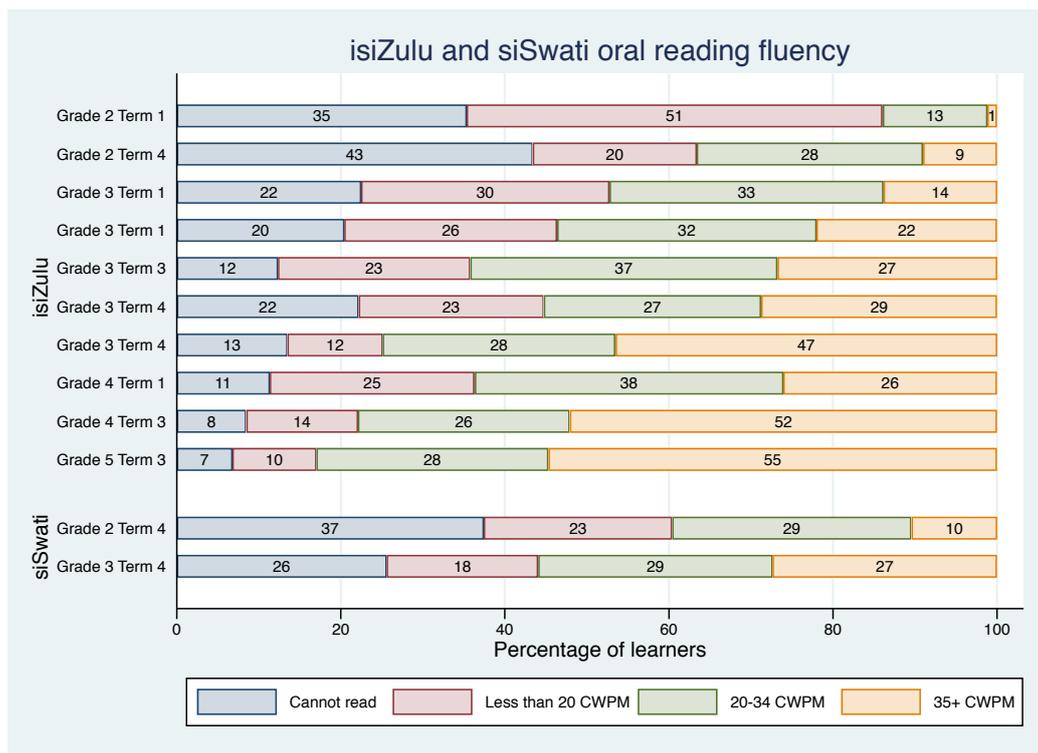
Next, we investigate whether our proposed thresholds correspond to meaningful and distinguishable zones along the reading development cline (see Figure 2). In Table 6, we combine all the samples within a language and summarise accuracy, comprehension and letter-sound knowledge by the same four fluency categories.

Table 6 shows that learners below the lower threshold (<20 cwpm):

♦ have low levels of accuracy and very poor comprehension scores;
♦ tend to perform poorly on letter-sound knowledge;
♦ would benefit from instruction on letter-sound knowledge (a foundational skill) and on improving their word decoding.

Among learners that meet the lower threshold (20 cwpm):

♦ levels of accuracy have improved with over 70% of these learners achieving at least 95% accuracy;
♦ comprehension skills are developing but remain poor.

Finally, the group of learners who meet the benchmark of 35 cwpm (upper threshold):

♦ are accurate readers;
♦ with comprehension scores in the range of 59% to 74%; [24]
♦ would benefit from instruction that focuses on improving their comprehension skills through increasing vocabulary and critical engagement with text.

---

[24] Note however, that skilled readers typically read with 95-100% comprehension. Readers who achieve 74% for reading comprehension benefit from direct comprehension instruction.

Table 6: Learner characteristics by early grade fluency profiles

	isiXhosa	isiZulu	siSwati
Cannot read: ORF = 0			
Mean correct letter-sounds per minute	12.8	10.8	18.3
% unable to sound one letter	12%	20%	10%
ORF: 1 - 19 cwpm			
Mean correct letter-sounds per minute	31.1	20	39.5
% with at least 95% accuracy	19%	25%	19%
Comprehension (% of total correct)	21%	18%	21%
Comprehension (% of attempted correct)	47%	51%	32%
ORF: 20 - 34 cwpm			
Mean correct letter-sounds per minute	47.8	29	53.2
% with at least 95% accuracy	71%	78%	76%
Comprehension (% of total correct)	46%	46%	53%
Comprehension (% of attempted correct)	65%	73%	62%
ORF: 35+ cwpm			
Mean correct letter-sounds per minute	57.2	36.3	61.8
% with at least 95% accuracy	87%	90%	84%
Comprehension (% of total correct)	59%	62%	74%
Comprehension (% of attempted correct)	73%	78%	74%

7.1.4.   Predictive validity of fluency thresholds

We used longitudinal data to investigate the extent to which meeting the proposed ORF thresholds predicts ORF in the second assessment. For example, using EGRS II data we followed siSwati learners from the last term of grade 2 to the last term of grade 3. We compared the grade 3 outcomes for learners who met and did not meet the thresholds at the end of grade 2[25]. Figures 12 to 14[26] categorise these learners' fluency at the second assessment based on three categories of learners' reading scores in assessment I:

i.   non-readers (0 cwpm) at assessment I (Figure 12);
ii.   those not meeting the lower threshold (1-19 cwpm) at assessment I (Figure 13);
iii. those meeting the lower threshold but not the upper threshold (20-34 cwpm) at assessment I (Figure 14).

Figure 12 shows that many learners who were unable to read at all at the first assessment were still unable to read one word by assessment II as shown by the blue bars. The percentage of these learners that are non-readers increases by grade, suggesting that chronic reading difficulties have set in among learners not able to read at all by grade 3. However, a sizeable portion of these non-readers have begun to read slowly by assessment II (3 to 6 terms later), but most are not yet reaching

---

25 For a summary of the timing of the first and second assessments, by sample, please see the technical report.
26 Within each group, the bar title indicates the grade and term of the first assessment.

the lower threshold of 20 cwpm (green bars). A very small percentage are reaching the upper threshold at the second assessment.

In Figure 13 we see that most learners (58% to 74%) who were reading between 1 and 19 cwpm at the first assessment, had reached the threshold of 20 cwpm at assessment II. Figure 14 shows that reaching the lower threshold by the first assessment is predictive of reaching the upper threshold by the time of the second assessment. Of these learners reaching the lower threshold by the first assessment, 51% to 78% were reading at or above the fluency benchmark of 35 cwpm by assessment II. This suggests that the proposed ORF thresholds are useful as they predict reading performance in later grades.

In addition to oral reading comprehension, two of the studies included a written comprehension exercise. We conducted a similar analysis comparing the relationship between oral reading fluency at the first assessment and performance on the written comprehension 4 to 6 school terms later. The results show that learners who had not met the fluency threshold at the first assessment performed poorly on the written comprehension at assessment II. For more detail, please see the technical report.

Figure 12: Reading fluency category at assessment II for non-readers at assessment I
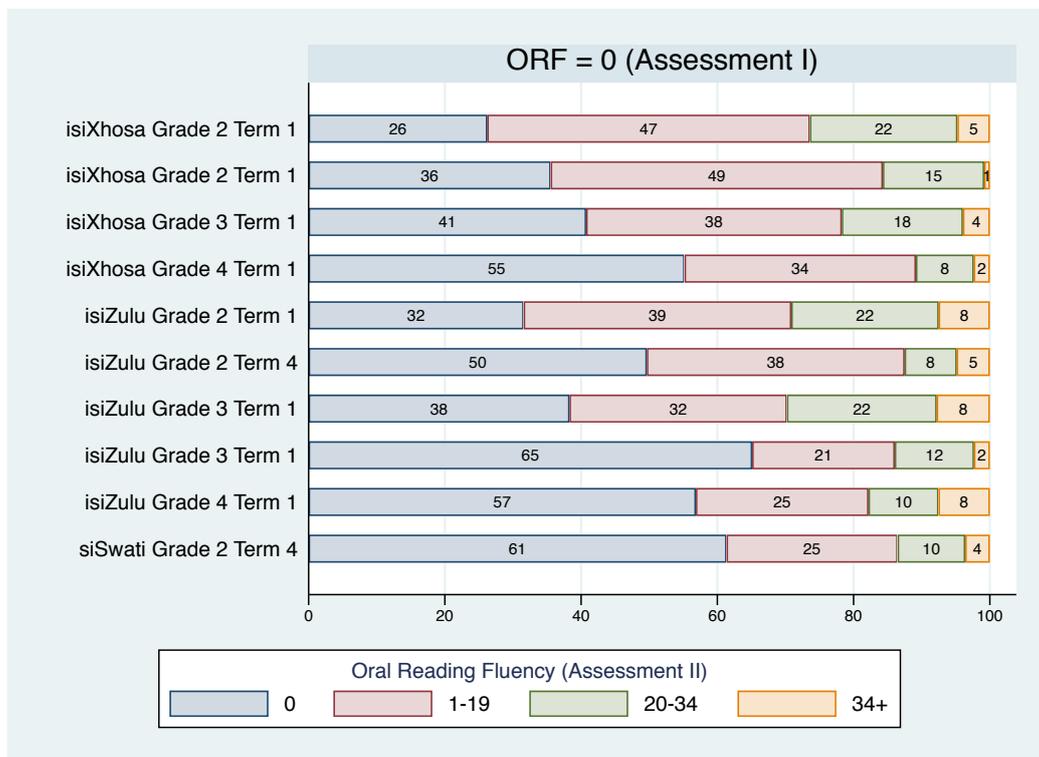
Figure 13: Reading fluency category at assessment II for learners scoring an ORF of 20-34 cwpm at assessment I

ORF = 1-19 (Assessment I)

	0	1-19	20-34	34+
isiXhosa Grade 2 Term 1	8	26	50	17
isiXhosa Grade 2 Term 1	3	38	51	7
isiXhosa Grade 3 Term 1	3	29	55	12
isiXhosa Grade 4 Term 1	4	29	63	4
isiZulu Grade 2 Term 1	5	21	53	21
isiZulu Grade 2 Term 4	2	30	55	13
isiZulu Grade 3 Term 1	5	25	47	23
isiZulu Grade 3 Term 1	1	27	57	15
isiZulu Grade 4 Term 1	4	30	50	16
siSwati Grade 2 Term 4	5	33	50	12

Oral Reading Fluency (Assessment II)
0   1-19   20-34   34+

Figure 14: Reading fluency category at assessment II for learners scoring an ORF of 1-19 cwpm at assessment I

ORF = 20-34 (Assessment I)

	0	1-19	20-34	34+
isiXhosa Grade 2 Term 1		2	32	66
isiXhosa Grade 2 Term 1		4	39	57
isiXhosa Grade 3 Term 1		2	35	62
isiXhosa Grade 4 Term 1	0	1	47	51
isiZulu Grade 2 Term 1	1	4	23	72
isiZulu Grade 2 Term 4		3	38	59
isiZulu Grade 3 Term 1	0		21	78
isiZulu Grade 3 Term 1	1	1	33	65
isiZulu Grade 4 Term 1	0	2	38	60
siSwati Grade 2 Term 4	1	3	45	51

Oral Reading Fluency (Assessment II)
0   1-19   20-34   34+

7.2. Establishing a letter-sound benchmark

Having identified a fluency threshold and benchmark, we now work backwards along the development cline to consider whether and what letter-sound benchmark could support the acquisition of fluency skills.

We first consider whether a letter-sound benchmark is a useful predictor of later fluency skills. We do this by using longitudinal data to explore the relationship between earlier letter-sound knowledge and later oral reading fluency. Figure 15 shows the relationship between correct letter-sounds per minute at the first assessment and oral reading fluency at the second assessment. Despite differences in the level and slope of the lines, there is a clear pattern: learners with better letter-sound knowledge at the first assessment have higher ORF scores when assessed again 3 to 6 terms later.

This suggests that reaching a letter-sound knowledge benchmark in the early grades would be a good indicator of whether a learner is on track to develop ORF in later grades. This provides a solid motivation for developing a letter-sound knowledge benchmark.

Figure 15: Relationship between letter-sound knowledge and later oral reading fluency
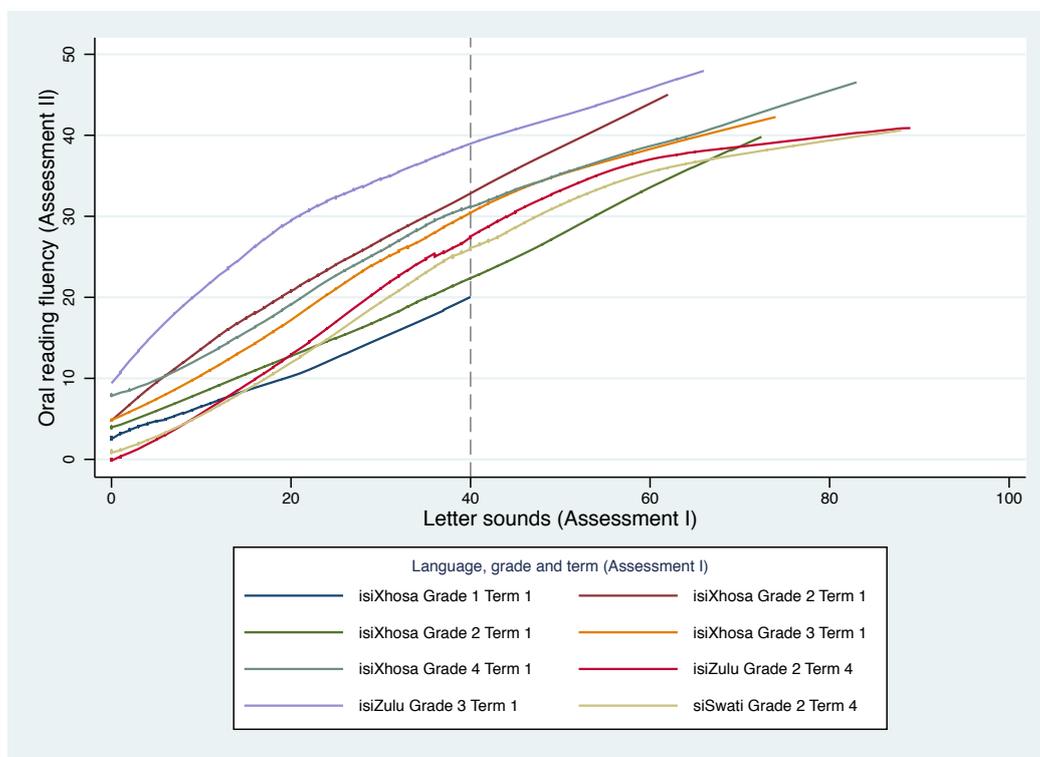
7.2.1. Relationship between speed and accuracy

But what would be an appropriate letter-sound benchmark? In answering this question, we again examine the relationship between speed and accuracy but now in relation to sounding letters. Figures 16 to 18 plot the relationship between the number of letter-sounds attempted (speed) and the percentage of these attempted letter-sounds that the learner reads correctly (accuracy) for each of the Nguni languages.

Figure 16: Relationship between letter-sound speed and accuracy, isiXhosa samples

isiXhosa letter sound speed and accuracy

Grade 1 Term 1	Grade 1 Term 4	Grade 2 Term 1	Grade 2 Term 1
Grade 2 Term 4	Grade 3 Term 1	Grade 3 Term 3	Grade 4 Term 1
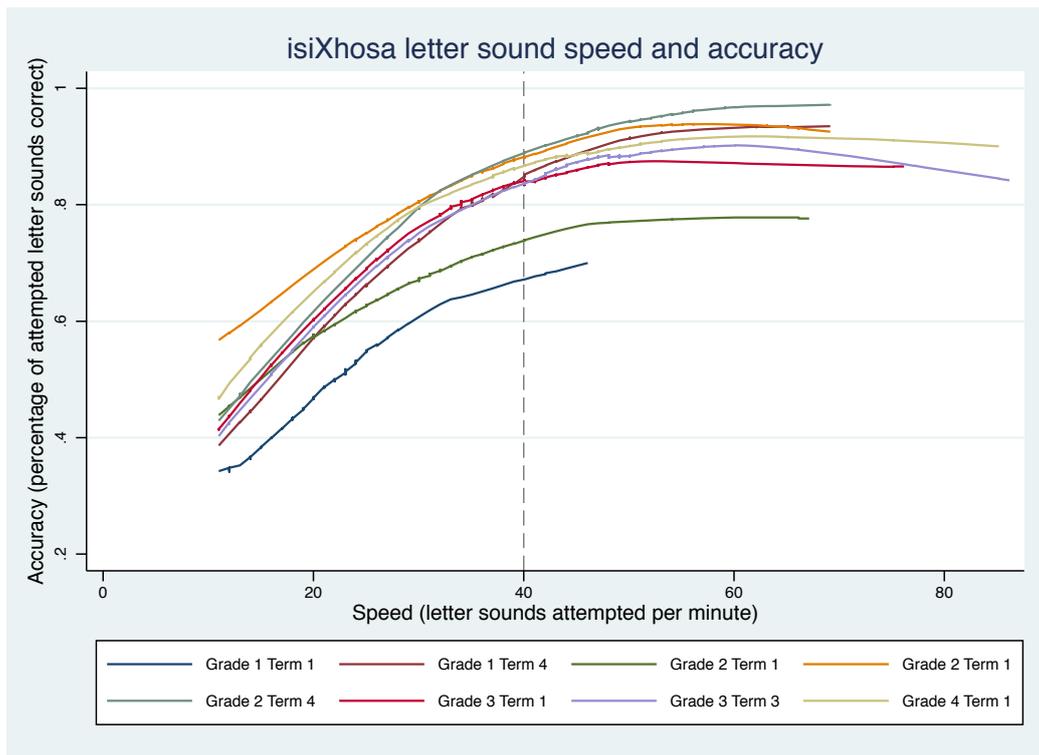
Figure 17: Relationship between letter-sound speed and accuracy, isiZulu samples

isiZulu letter sound speed and accuracy

|	
Grade 2 Term 4	Grade 3 Term 1
Grade 3 Term 4	Grade 3 Term 4
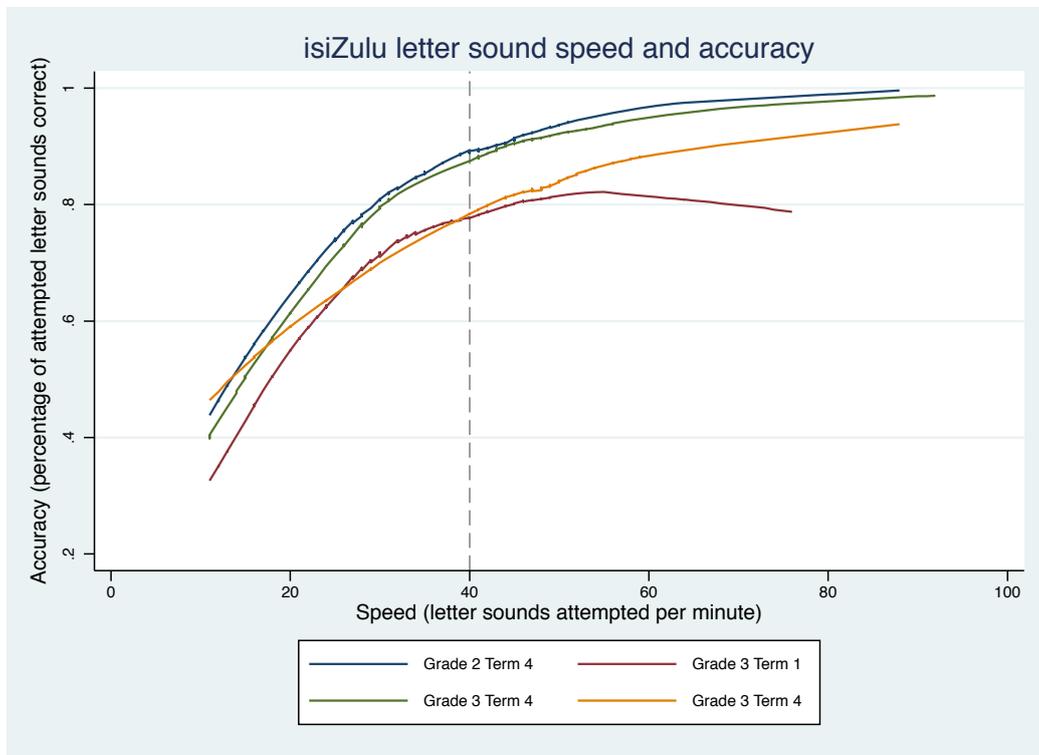
Figure 18: Relationship between letter-sound speed and accuracy, siSwati samples
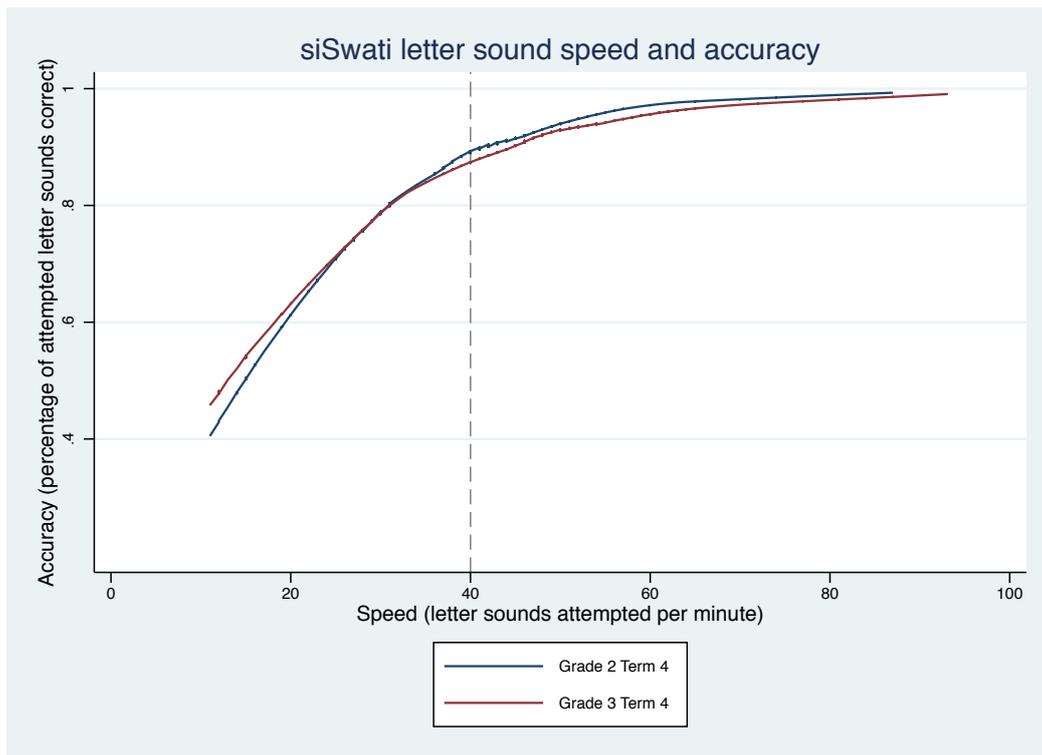
The figures confirm the following patterns with respect to reading letter-sounds:

♦ Learners with low speed tend to have low accuracy.
♦ Accuracy improves steadily with speed to a point, but beyond this point there are no further improvements in accuracy.
♦ The letter-sound speed-accuracy gradient tends to flatten around 40 letter-sounds per minute (indicated by the dashed grey line).

Analysis of the longitudinal data, shown in the technical report, provides support for the notion that there is letter-sound recognition speed beyond which learners show little to no improvement. These patterns suggest that a letter-sound benchmark of around 40 correct letter-sounds per minute is appropriate.

7.2.2. Predictive validity of the proposed letter-sound benchmark

We examined how the proposed benchmark of 40 letter-sounds cpm maps to future oral reading fluency in Figures 19 to 21. Figure 19 shows fluency at assessment II for learners who met the letter-sound benchmark at the first assessment. Figure 20 shows fluency at assessment II among learners who were correctly sounding out 20 to 39 letters cpm at the first assessment. Figure 21 shows learners who had low or no letter-sound proficiency when first assessed.

Figure 19: Reading fluency at assessment II, for learners who achieved the proposed letter-sound benchmark at assessment I

Letter Sounds = 40+ (Assessment I)

	0	1-19	20-34	35+
isiXhosa Grade 1 Term 1			83	17
isiXhosa Grade 2 Term 1		8	45	46
isiXhosa Grade 2 Term 1	2	17	51	30
isiXhosa Grade 3 Term 1	1	8	42	49
isiXhosa Grade 4 Term 1	0	5	43	51
isiZulu Grade 2 Term 4	1	6	37	56
isiZulu Grade 3 Term 1	1	3	25	71
siSwati Grade 2 Term 4	2	9	40	49

Percentage of learners

Oral Reading Fluency (Assessment II)
0   1-19   20-34   35+
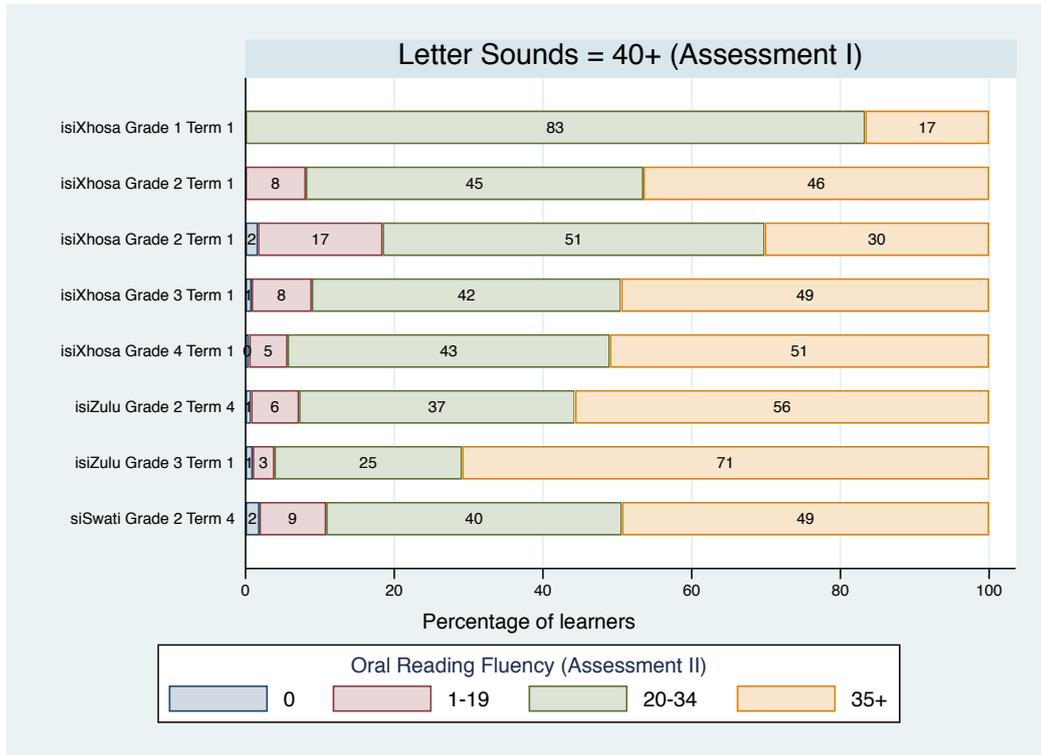
Figure 20: Reading fluency at assessment II, for learners who could read 20 to 39 letter-sounds correctly at assessment I

Letter Sounds = 20-39 (Assessment I)

|	0	1-19	20-34	35+
isiXhosa Grade 1 Term 1	12	59	29	
isiXhosa Grade 2 Term 1	1	25	53	21
isiXhosa Grade 2 Term 1	7	49	37	7
isiXhosa Grade 3 Term 1	7	29	44	20
isiXhosa Grade 4 Term 1	4	21	48	27
isiZulu Grade 2 Term 4	13	35	31	21
isiZulu Grade 3 Term 1	6	11	32	51
siSwati Grade 2 Term 4	19	30	33	18

Oral Reading Fluency (Assessment II)
0   1-19   20-34   35+
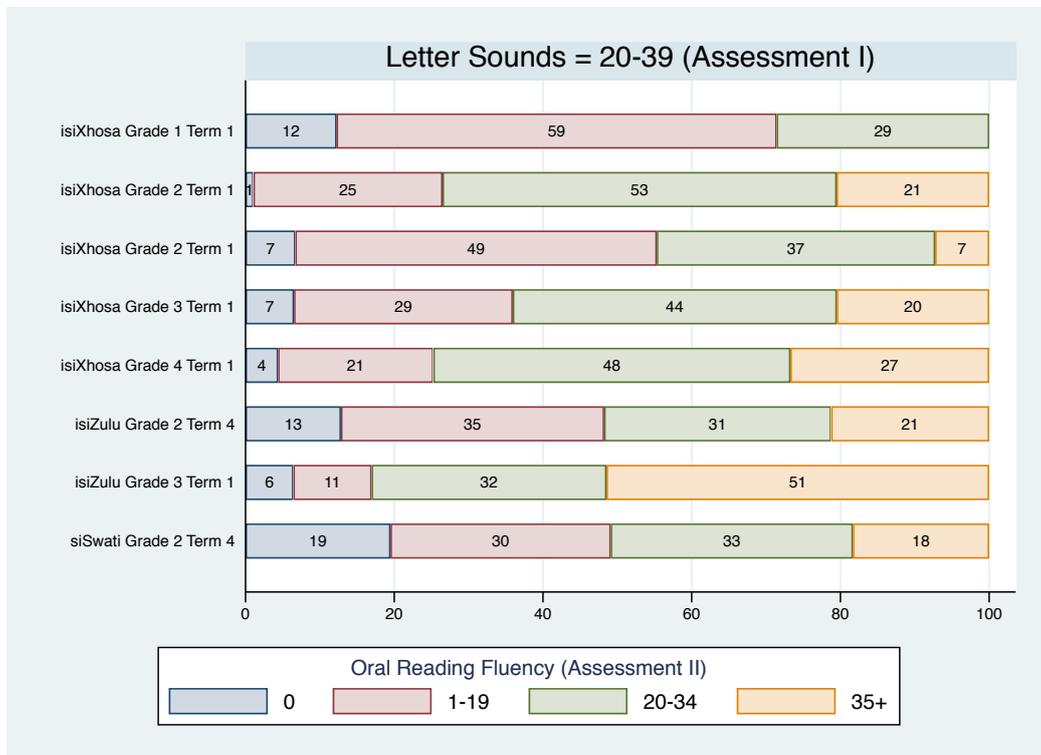
Figure 21: Reading fluency at assessment II, for learners who sounded less than 20 letter-sounds correctly at assessment I

Letter Sounds = 0-19 (Assessment I)

Category	0	1-19	20-34	35+
isiXhosa Grade 1 Term 1	60	32	8	0
isiXhosa Grade 2 Term 1	27	44	24	6
isiXhosa Grade 2 Term 1	38	47	14	0
isiXhosa Grade 3 Term 1	38	34	22	6
isiXhosa Grade 4 Term 1	31	24	32	14
isiZulu Grade 2 Term 4	61	30	6	3
isiZulu Grade 3 Term 1	30	19	26	25
siSwati Grade 2 Term 4	69	22	7	2

Oral Reading Fluency (Assessment II): 0 | 1-19 | 20-34 | 35+
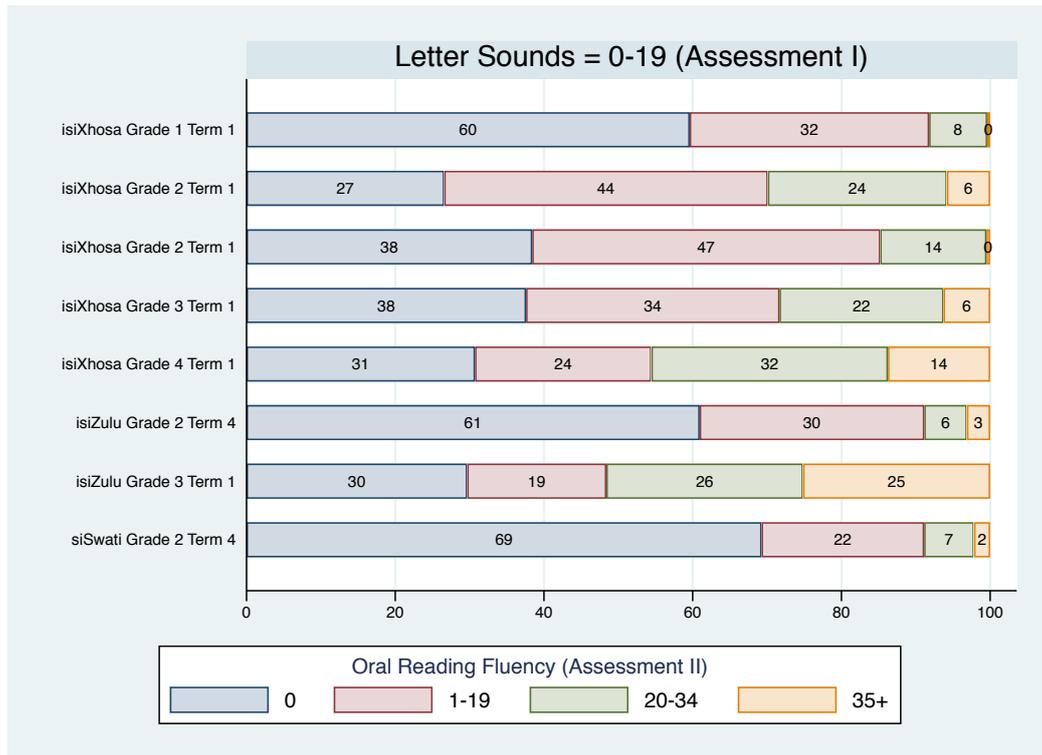
Among learners who met the proposed letter-sound benchmark of 40 at assessment I, many had reached the fluency benchmark of 35 cwpm when they were re-assessed 3 to 6 school terms later. If they missed the benchmarks of 35 cwpm, most had at least reached the lower threshold of 20 cwpm.

In most samples, among learners who were achieving 20 to 39 letter-sounds cpm at the first assessment, significant proportions met the lower fluency threshold of 20 cwpm by the second assessment.

Among learners who had low letter-sound proficiency (0 to 19 letter-sounds cpm) when first assessed, a sizeable portion were unable to read one word at the next assessment. Very few reached the fluency benchmark of 35 cwpm when re-assessed.

In summary, the letter-sound benchmark of 40 letters sounded correctly per minute appears to be a good predictor of oral reading fluency in later grades.

8. Summary and recommendations

This research aimed to establish early grade reading benchmarks in three Nguni languages. Through an exploratory analysis of available data, positioned within a theoretical framework that outlines the reading development process, the report identifies a letter-sound benchmark, an oral reading fluency (ORF) threshold and an ORF benchmark to guide the teaching of reading in the Foundation Phase. Given the close relationship between learner proficiency in single word reading and ORF in the available data, it was unnecessary to set a word reading benchmark.

As a preliminary step in the analysis, current reading norms in Nguni languages for predominately rural-based learners in no-fee schools were identified from available EGRA data sources. By the end of grade 3, this learner sample read an average of 19 to 25 cwpm from a passage, across the three languages. The proportion of learners that are non-readers by the end of grade 3 remains unacceptably high (15% to 26% of Nguni-speaking learners in these data samples were unable to read one word from a passage).

8.1. Summary of proposed benchmarks and thresholds

Based on the exploratory data analysis presented in this report, we propose the following threshold and benchmarks for early reading success in Nguni languages.

- ♦ A letter-sounds benchmark of 40 letter-sounds correct per minute (cpm), which should be reached by all learners at the end of grade 1.
- ♦ A lower ORF threshold of 20 cwpm from a connected text. All learners should reach this minimum threshold by the end of grade 2.
- ♦ An ORF benchmark of 35 cwpm, which should be attained by all learners by the end of grade 3.

Learners who reach the letter-sounds benchmark of 40 letters cpm are in a much stronger position to read words in and out of context, and decoding instruction can focus on developing fluency. There are few remaining benefits to improving letter-sound knowledge and speed beyond this point. Meeting this letter-sound benchmark is shown to be an early predictor of whether a learner is on track to reach our recommended ORF threshold and benchmark in the later grades.

The ORF threshold of 20 cwpm is a *minimum* threshold which learners must reach so that instructional focus can facilitate the engagement of higher order reading skills during reading. ***Below this threshold we find little evidence that learners can comprehend what they have read, evidenced by very low oral comprehension or written comprehension scores.*** Attention to the development of fluency and prosody should continue so that reading aloud sounds increasingly natural.

The ORF benchmark of 35 cwpm acts as an upper threshold. Once learners reach the upper threshold of 35 cwpm when reading connected text, attention to prosody continues but the main instructional focus should be on developing reading comprehension strategies and improving vocabulary, to empower learners to engage critically with texts. Learners who reach this upper benchmark will be able to transition to silent reading more easily.

Currently, by the end of grade 3, most learners (53% to 76%, depending on the sample) have only reached the lower threshold of 20 cwpm. Approximately a quarter of learners have reached the fluency benchmark (35 cwpm) and would benefit from a strong instructional focus on comprehension skills.

In addition to being empirically and theoretically grounded, we believe the suggested benchmarks

and threshold are valid for the South African context for three main reasons:

i.   They are neither too aspirational nor too low. The suggested benchmarks and threshold are not so high as to be completely out of reach for current Foundation Phase learners in poor South African school contexts - a fact that we show in the data. Yet, they establish expectations that are also sufficiently high enough to move teachers, classrooms and schools on a journey of reading excellence, aligning with the presidential goal that all learners should be able to read for meaning by the end of grade 3.

ii. Strong theoretical foundations and predictive validity: They meet other criteria for a "good" threshold as they are theoretically sound, and they predict reading performance in later grades.

iii. Importantly, as explained in the methodology section, we make no assumptions about the fluency-comprehension relationship in the context of South African classrooms and in the Nguni languages. Rather, we allow the data to guide us in the selection of appropriate thresholds and benchmarks. This process is independent of the difficulty level of comprehension questions. A weakness of other benchmarking approaches is that, while the analysis is easier to perform, their results are highly sensitive to the difficulty and placement of comprehension questions in the reading assessment.

8.2. Policy recommendations

The Nguni language benchmarks and threshold proposed in this report provide standards to establish whether learners are on a healthy reading trajectory; guiding teachers to identify learners who are at risk of falling behind in the developmental sequence of reading. In particular, the letter-sound threshold provides a tool for early identification of at-risk learners which can be easily and quickly administered. The proposed benchmarks and threshold could be used to guide classroom assessment of letter-sound knowledge and oral reading assessment.

The very high number of learners entering grade 1 with no letter-sound knowledge suggests that the quality of grade R instruction must be evaluated. From the limited data available, this report also highlighted that learners find complex consonant sequences much more difficult than reading and sounding single letter-sounds. Because complex sounds occur frequently in the Nguni languages, knowledge of complex consonants is required by the end of grade 1 to enable learners to keep on track with the proposed ORF threshold and benchmark by grade 2 and 3. Resolving this would require adaptation of the curriculum which guides Nguni language instruction. Currently the teaching of complex consonant sequences is not included in the grade 1 curriculum.

In closing we return to the ultimate goal of reading: *reading for meaning*. Our EGRA sample data confirms patterns observed in national PIRLS testing of significant deficits in learners' acquisition of comprehension skills. While the teaching of foundational reading skills is fundamental to reading proficiently, the threshold and benchmarks proposed are a necessary but not sufficient condition for reading for meaning. In other words, even the ORF benchmark of 35 cwpm does not represent the point where learners are able to read for meaning on their own. However, at this level of fluency, comprehension is a possibility, if the skills to answer oral and written comprehension are taught. It is evident from our analysis that more attention needs to be given in the Foundation Phase to the teaching of comprehension skills. But the teaching of comprehension skills, and higher order skills may have to be differentiated, depending on where learners are on the developmental sequence of reading. While it may be a challenge to provide learner-specific guidance in Foundation Phase classroom contexts with large class sizes, group-guided reading provides the opportunity for differentiated learner strategies to meet each learner along their reading development path. We provide another instructive signal to support this: For those learners who start approaching the ORF benchmark of 35 cwpm, this is the signal to boost the teaching of reading comprehension skills.

9. References

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*. Stellenbosch: ReSEP, Stellenbosch University. Cape Town: SALDRU, University of Cape Town. Chicago: NORC at the University of Chicago. Pretoria: Department of Basic Education

Ardington, C. & Menendez, A. (2020). *Comments on Design Report B: Reading Benchmarking in South Africa*. Unpublished report submitted to USAID. Cape Town.

Ardington, C., Hoadley, U. and Menendez, A. (2019). *Impact Evaluation of USAID/South Africa Story Powered School Program Endline Report.* Washington, DC: USAID Reading and Access Activity.

Caravolas, M., Lervåg, A., Defior, S., Seidlova-Makora, G. & Hulme, C. (2013). Different Patterns, but Equivalent Predictors, of Growth in Reading in Consistent and Inconsistent Orthographies. *Psychological Science*, 24(8), 1398-1407.

Department of Basic Education, (2019). *The Secondary Early Grade Reading Study. Year 1 Report. Learner performance after the first year of implementation*. DBE, Pretoria.

Fuchs, L.S., Fuchs, D., Hosp, M.K. & Jenkins, J.R. (2001) *Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis*. Scientific Studies of Reading, 5(3), 239–256.

Fuchs, D., Compton, D., Fuchs, L., Bryant, J., Hamlett, C. & Lambert, W. (2012). First-grade cognitive abilities as long-term predictors of reading comprehension and disability status. *Journal of Learning Disabilities*, 45, 217–231. doi:10.1177/0022219412442154

García, J.R. & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84, 74–111. doi:10.3102/0034654313499616

Gough, P. & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6-10.

Hasbrouck, J. & Tindal, G.A. (2006). Oral reading fluency norms: a valuable assessment tool for reading teachers. *The Reading Teacher*, 59(7), 636-644.

Hoover, W.A. & Gough, P.B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127–160. doi:10.1007/BF00401799

Howie, S., Combrinck, C., Roux, K., Tshele, M., Mokoena, G., & McLeod Palane, N. (2017). *Progress in International Reading Literacy Study 2016 - South African Childrens' Reading Literacy Achievement. Pretoria*: Centre for Evaluation and Assessment

Jukes. M., Cummiskey, C., & Gargano, M. (2018). *Guidance Note: Setting Data-driven Oral Reading Fluency Benchmarks*. North Carolina: RTI International

Pikulski, J.J. & Chard, D.J. (2006). Bridge between decoding and reading comprehension. *The Reading Teacher*, 58(6), 510–519. doi:10.1598/RT.58.6.2
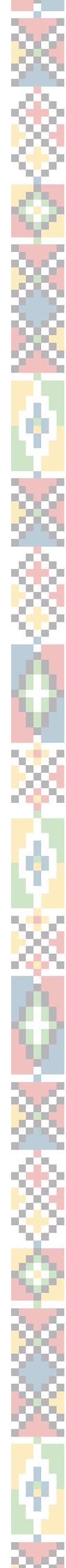
RTI International (2017). *All Children Reading-Asia: EGRA Benchmarks and Standards Research Report*. https://shared.rti.org/content/all-children-reading-asia-egra-benchmarks-and-standards-research-report

South African Government (2019). "State of the Nation Address by Cyril Ramaphosa, 2019." June 20. https://www.gov.za/speeches/2SONA2019.

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 for Early Childhood Education* 10(1), a773. doi:10.4102/sajce.v10i1.773

Spear-Swerling, L. (2006) Children's reading comprehension and oral reading fluency in easy text. Reading and Writing: An Interdisciplinary Journal, 19, 199-220.

Taylor, N., Wills, G. and Hoadley, H. (2019). Addressing the 'leadership conundrum' through a mixed methods study of school leadership for literacy. *Research in Comparative and International Education* 14 (1), 30-53. doi: 10.1177/1745499919828928.

Torppa, M., Georgiou, G.K., Lerkkanen, M.K., Niemi, P., Poikkeus, A.M., & Nurmi, J.E. (2016). Examining the simple view of reading in a transparent orthography: A longitudinal study from kindergarten to grade 3. *Merrill Palmer Quarterly*, 62, 179–206. http://digitalcommons.wayne.edu/mpq/vol62/iss2/4/

Tunmer, W.E. & Chapman, J.W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, 45, 453– 466. doi: 10.1177/0022219411432685

Wang, Z, Sabatini, J, O'Reilly, T. & Weeks, J. (2019). Decoding and Reading Comprehension: A Test of the Decoding Threshold Hypothesis. *Journal of Educational Psychology*, 111(3), 387-401.

Zenex Foundation (2018). *Leveraging our collective wisdom to address the literacy challenge*. Conference Proceedings of the Literacy Symposium, 24 October 2018. https://www.zenexfoundation.org.za/informationhub/publications/early-grades-numeracy-literacy

Funders of this report