



BENCHMARKING EARLY GRADE READING SKILLS IN SOUTH AFRICA: TECHNICAL REPORT

ENGLISH FIRST ADDITIONAL LANGUAGE

15 June 2022

This publication was produced at the request of the United States Agency for International Development. It was prepared independently by Khulisa Management Services, (Pty) Ltd in collaboration with the South African Department of Basic Education

Photo: Teaching early grade literacy, Mogokonyane Primary School, North West Province, South Africa

Photo credit: Khulisa Management Services

AUTHORS

Dr Gabrielle Wills (Research On Socio-Economic Policy, Stellenbosch University)

Prof Cally Ardington (University Of Cape Town, SALRDU)

Prof Elizabeth Pretorius (Emerita Professor And Research Fellow In The Department Of Linguistics And Modern Languages, Unisa)

Dr Eileen Pooe (Senior Lecturer, North West University & Setswana National Language Body - PanSALB)

Dr Refilwe Ramagoshi (Research Associate and Senior Lecturer, University of Pretoria)

Ms Lesang Sebaeng (Department of Basic Education)

ACKNOWLEDGEMENTS

Thank you to Ms Nompumelelo Mohohlwane and Ms Lesang Sebaeng from the Department of Basic Education for their valuable contributions to the development of the Learner Assessment passages and tools, training of fieldworkers and quality assurance of the data. In addition, we would like to thank the North West Province officials, school principals, teachers and learners for participating in this study. This study would not have been possible without the commitment of our partners and the Khulisa team, including the fieldworkers and supervisors who spent many hours collecting the data, and the Education Specialist, Ms Benita Williams.

CONTACT DETAILS

Margaret Roper
Khulisa Management Services
26 7th Avenue
Parktown North
Johannesburg, 2196
Telephone: 011-447-6464
Email: mroper@khulisa.com
Web Address: www.khulisa.com

DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Abbreviations and Acronyms

CAPS	Curriculum and Assessment Policy Statement
clspm	correct letter sounds per minute
cwpm	correct words per minute
DBE	Department of Basic Education
DIBELS	Dynamic Indicators of Basic Early Literacy Skills
EC	Eastern Cape
EFAL	English First Additional Language
EGRA	Early Grade Reading Assessment
EGRS	Early Grade Reading Study
ESL	English Second Language
FW	Funda Wandé
GoSA	Government of South Africa
HL	home language
KZN	KwaZulu-Natal
L1	First language
L2	Second language
LFL	Leadership for Literacy
LOLT	language of learning and teaching
LP	Limpopo Province
LSK	letter sound knowledge
MP	Mpumalanga Province
NEEDU	National Education Evaluation and Development Unit
NW	North West
ODH	orthographic depth hypothesis
ORF	oral reading fluency
PGST	psycholinguistic grain size theory
PIRLS	Progress in International Reading and Literacy Study
RC	reading comprehension
RSP	Reading Support Project
SPS	Story Powered Schools
SVR	simple view of reading
USA	United States of America

Table of Contents

PREAMBLE	i
EXECUTIVE SUMMARY	ii
I INTRODUCTION	1
1.1 Background	2
1.1.1 Aims	4
1.1.2 Report structure.....	6
2 THEORETICAL FRAMEWORK FOR EARLY READING DEVELOPMENT	6
2.1 Reading theory: Three relevant models	7
2.1.1 Simple view of reading	7
2.1.2 The decoding threshold hypothesis	8
2.1.3 Orthographic depth hypotheses	9
2.1.4 Instructional practice shapes reading development.....	10
2.2 A developmental view of multiple proficiencies	11
3 BENCHMARKING METHODOLOGY	13
3.1 Approach	13
3.2 Statistical method and skills	14
3.2.1 Exploratory non-parametric techniques	14
3.2.2 Advantages of non-parametric methods over other benchmarking approaches	14
3.2.3 Establishing ORF benchmarks.....	15
3.2.4 Establishing letter sound benchmarks	16
3.3 Instrument development process: Instruments that support benchmarking	16
4 EARLY READING DEVELOPMENT IN ENGLISH SECOND LANGUAGE (ESL) OR FIRST ADDITIONAL LANGUAGE (EFAL)	19
4.1 International evidence on norms and benchmarks in English reading.....	19
4.2 Evidence from Southern Africa	20
4.2.1 Reading levels in English	20
5 BENCHMARKING METHODOLOGY: EFAL.....	25
6 EFAL DATA.....	26
6.1 EFAL ORF passage difficulty.....	28
6.2 EFAL school sample characteristics	29
6.3 EFAL reading norms in no-fee schools in six provinces: Oral reading fluency and comprehension	30
7 BENCHMARKING RESULTS: EFAL	34
7.1 Establishing a fluency threshold and benchmark in EFAL: An analysis of reading speed and accuracy	34
7.2 Fluency and comprehension.....	38
7.3 Predictive validity of EFAL fluency threshold and benchmark	43

8	ALIGNING EFAL FLUENCY THRESHOLDS AND BENCHMARKS TO GRADE LEVELS	46
8.1	Examining the attainability of the lower threshold and benchmark.....	46
8.2	Examining the lower threshold and benchmark against grade level English fluency norms from other contexts.....	48
8.3	Establishing a Grade 2 development fluency threshold in EFAL	51
9	SUMMARY: EFAL FLUENCY THRESHOLD AND BENCHMARKS.....	52
10	RECOMMENDATIONS	55
	REFERENCES	58

List of Tables

TABLE E 1: EFAL FLUENCY LEVELS AT THE END OF EACH PRIMARY SCHOOL GRADE AMONG NO-FEE LEARNER SAMPLES FROM 6 PROVINCES (EGRA-TYPE STUDIES FROM 2017-2021).....	IX
TABLE 1: DESCRIPTIONS OF SOUTHERN AFRICAN STUDIES IDENTIFYING LETTER-SOUND KNOWLEDGE AND ORAL READING FLUENCY IN ENGLISH AMONG EARLY GRADE READERS.....	21
TABLE 2: AVERAGE LETTER-SOUND KNOWLEDGE AND ORAL READING FLUENCY IN ENGLISH AMONG EARLY GRADE READERS, EVIDENCE FROM SOUTHERN AFRICAN STUDIES.....	24
TABLE 3: AVAILABLE SAMPLES BY YEAR, TERM, AND GRADE TO EXPLORE ORAL READING FLUENCY IN EFAL	27
TABLE 4: ORF COMPREHENSION PERFORMANCE FOR LEARNERS WITH ORF > 0 AND ALLOWED TO READ THE ORF PASSAGE FOR 3-MINUTES.....	28
TABLE 5: READABILITY OF SELECT EFAL PASSAGES.....	29
TABLE 6: SCHOOL CHARACTERISTICS OF THE EFAL DATA SAMPLE.....	30
TABLE 7: WEEKLY TIME ALLOCATION TO FIRST ADDITIONAL LANGUAGE VS. HOME LANGUAGE IN CAPS.....	32
TABLE 8: ORF AND ORF COMPREHENSION SCORES, SOUTH AFRICAN EARLY GRADE READING ASSESSMENTS (2017-2021).....	33
TABLE 9: MEDIAN EFAL FLUENCY (50 TH PERCENTILE), SOUTH AFRICAN LEARNER SAMPLES (INCLUDING REPEATERS).....	48
TABLE 10: ENGLISH FLUENCY NORMS FROM INTERNATIONAL STUDIES - END OF GRADE MEDIAN FLUENCY (50 TH PERCENTILE) OR IN THE CASE OF DIBELS THE ORF ASSOCIATED WITH BEING AT NEGLIGIBLE OR MINIMUM RISK OF READING FAILURE AT THE END OF EACH GRADE.....	49
TABLE 11: CURRENT STATUS OF DBE DONOR/PARTNER PLANNED EARLY GRADE READING BENCHMARKING IN ALL OFFICIAL LANGUAGES.....	56
TABLE A 1. PILOT SAMPLES	63
TABLE A 2: MORE DETAILS ON SOUTH AFRICAN EGRA-TYPE STUDIES USED TO ASSESS READING SUBSKILLS IN EFAL.....	64

List of Figures

FIGURE E 1: THREE RESEARCH COMPONENTS INTERSECT TO ESTABLISH EARLY GRADE READING THRESHOLDS AND BENCHMARKS.....	V
FIGURE E 3: FLUENCY BENCHMARKS AND THRESHOLDS FOR EARLY GRADE READING IN EFAL	VIII
FIGURE 1: DEVELOPMENTAL CLINE IN EARLY READING	11
FIGURE 2: DISTRIBUTION ORAL READING FLUENCY SCORES (CVPM) FOR RURAL SOUTH AFRICAN ESL LEARNERS RELATIVE TO BROWARD COUNTY ESL LEARNERS (FROM DRAPER AND SPAULL, 2015).....	23
FIGURE 3: PERCENTAGE OF LEARNERS SCORING ZERO ON ORF ASSESSMENTS IN EFAL, SOUTH AFRICAN EARLY GRADE READING ASSESSMENTS (2017-2021)	34
FIGURE 4: READING SPEED AND ACCURACY IN EFAL, BY SAMPLE.....	35
FIGURE 5: SPEED DISTRIBUTION FOR LEARNERS READING WITH AT LEAST 85% ACCURACY IN EFAL	37
FIGURE 6: SPEED DISTRIBUTION FOR LEARNERS READING WITH AT LESS THAN 85% ACCURACY IN EFAL	37
FIGURE 7: RELATIONSHIP BETWEEN FLUENCY AND COMPREHENSION FOR SAMPLES COMPLETING A SUB-SET OF ALL COMPREHENSION QUESTIONS, EFAL.....	38
FIGURE 8: GRADE 3 EFAL ORF DISTRIBUTION FOR A GIVEN WRITTEN COMPREHENSION RAW SCORE (EGRS I 2018)	39
FIGURE 9: GRADE 4 EFAL ORF DISTRIBUTION FOR A GIVEN WRITTEN COMPREHENSION RAW SCORE (EGRS II 2020)	40
FIGURE 10: EFAL ORF DISTRIBUTION FOR A GIVEN ORF COMPREHENSION RAW SCORE	41
FIGURE 11: GRADE 5 EFAL ORF DISTRIBUTION FOR A GIVEN WRITTEN COMPREHENSION RAW SCORE (SPS 2019)	42
FIGURE 12: GRADE 6 EFAL ORF DISTRIBUTION FOR A GIVEN WRITTEN COMPREHENSION RAW SCORE (FW LP, 2021)	42
FIGURE 13: GRADE 7 EFAL ORF DISTRIBUTION FOR A GIVEN WRITTEN COMPREHENSION RAW SCORE (EGRS I, 2021)	43
FIGURE 14: FLUENCY IN GRADE 5, BY LEARNERS' FLUENCY PROFILE IN GRADE 4, EFAL	44
FIGURE 15: FLUENCY IN GRADE 7, BY LEARNERS' FLUENCY PROFILE IN GRADE 4, EFAL	45
FIGURE 16: FLUENCY IN GRADE 4, BY LEARNERS' FLUENCY PROFILE IN GRADE 3, EFAL	45
FIGURE 17: FLUENCY PROFILES IN EFAL BY GRADE (EXCLUDING REPEATERS).....	47
FIGURE 18: GRADE 2 FLUENCY BY LEARNER'S GRADE 3 FLUENCY CATEGORY A YEAR LATER, EGRS II (MPUMALANGA)	52
FIGURE A 1: RELATIONSHIP BETWEEN EFAL FLUENCY AND INDIVIDUAL COMPREHENSION QUESTIONS: RSP GRADE 4, TERM 3	65
FIGURE A 2: RELATIONSHIP BETWEEN EFAL FLUENCY AND INDIVIDUAL COMPREHENSION QUESTIONS: EGRS I GRADE 3, TERM 3	66
FIGURE A 3: RELATIONSHIP BETWEEN EFAL FLUENCY AND INDIVIDUAL COMPREHENSION QUESTIONS: EGRS II GRADE 4, TERM 4	67
FIGURE A 4: RELATIONSHIP BETWEEN EFAL FLUENCY AND INDIVIDUAL COMPREHENSION QUESTIONS: SPS II GRADE 4, TERM 3	68
FIGURE A 5: RELATIONSHIP BETWEEN FLUENCY AND INDIVIDUAL COMPREHENSION QUESTIONS: SPS II GRADE 5, TERM 3, EFAL.....	69
FIGURE A 6: FLUENCY IN GRADE 3, BY LEARNERS' FLUENCY PROFILE IN GRADE 2, EFAL.....	70

PREAMBLE

This report, produced by Khulisa Management Services (Pty) Ltd. (Khulisa), is submitted under the Data Collection and Analysis for the Early Grade Reading Study (EGRS), the Reading Support Project (RSP) and Language Benchmarking to the United States Agency for International Development (USAID) under PERFORMANCE Indefinite Delivery Indefinite Quantity (IDIQ) Contract Number: 72067418D00001, Order Number: 72067421F00001.

This report derives from the 2021 data collection and analysis for the EGRS I (wave 5 data), the RSP Impact Evaluations and the Language Benchmarking study in two districts in North West Province, South Africa.

A number of reports have been published under this task order and are useful as background.

- Methodology Plan and Study Protocol: Data Collection and Analysis for the EGRS, RSP and Benchmarking. https://pdf.usaid.gov/pdf_docs/42132810ec2c48809efe8ca11e155aff.pdf
- For the full instrument development process refer to the “Report on the Development of Learner Assessment Tools and Contextual Tools” <https://dec.usaid.gov/dec/home/Default.aspx>.
- The Quality Assurance Surveillance Protocol (QASP). The QASP documents the quality assurance elements of both data collection and analysis. https://pdf.usaid.gov/pdf_docs/PA00Z8SX.pdf
- Task Order 4 Data Collection and Analysis EGRS, RSP, Benchmark and COVID-19: Fieldwork Report https://pdf.usaid.gov/pdf_docs/e4563ed819164a79956698c3a1998964.pdf

As part of this task order, Khulisa conducted additional research on COVID-19 in the schools and two reports were provided. The Preliminary COVID-19 Report submitted in 2021 enabled the DBE to consider the policy implications to prepare for the 2022 school year. https://pdf.usaid.gov/pdf_docs/PA00XGST.pdf. Thereafter, the Consolidated Final COVID-19 Report was submitted in 2022 https://pdf.usaid.gov/pdf_docs/PA00ZBHD.pdf.

The EGRS I Impact Evaluation report and further reports on the EGRS are available on the Department of Basic Education website <https://www.education.gov.za/Programmes/EarlyGradeReadingStudy.aspx>.

Data was analysed to recommend Setswana Home Language (HL) reading benchmarks and English First Additional Language (EFAL) reading benchmarks. The complete technical reports for EFAL and Setswana HL, as well as Summary Reports and Learning Briefs are available on the USAID Development Experience Clearinghouse and the Department of Basic Education website <https://www.education.gov.za/Programmes/EarlyGradeReadingStudy.aspx>

The methodology for Setting Reading Benchmarks In South Africa is outlined in this report https://pdf.usaid.gov/pdf_docs/PA00X1NZ.pdf.

The data used for EFAL was based on studies funded by the Department of Basic Education, USAID, Zenex Foundation, UNICEF, the Allan Gray Orbis Foundation Endowment, and the Economic and Social Research Council.

EXECUTIVE SUMMARY

This report, produced by Khulisa Management Services (Pty) Ltd. (Khulisa), is submitted under the Data Collection and Analysis for the Early Grade Reading Study (EGRS), the Reading Support Project (RSP) and Benchmarking to the United States Agency for International Development (USAID) under Indefinite Delivery Indefinite Quantity (IDIQ) Contract Number: 72067418D00001, Order Number: 72067421F00001.

This report derives from the data collection and analysis for the EGRS I, the RSP and the Language Benchmarking study in two districts (Dr Kenneth Kaunda and Ngaka Modiri Molema) in North West, South Africa. Multiple assessment points for almost 16,000 Setswana home language (HL), who are English first additional language (EFAL) learners from 230 quintile 1-3 or no-fee schools¹ in North West Province are available from existing and new data collected through the first Early Grade Reading Study (EGRS I)² and Reading Support Project (RSP)³.

Reading for meaning is a core skill that children must master in the early grades to be able to learn. The South African school curriculum develops from the assumption that children can read for meaning in both their home language and English by the end of the Foundation Phase (Grade 3).

Reiterating the importance of learning to read to be able to learn, South African President Cyril Ramaphosa articulated a clear goal for basic education: every child should be able to read for meaning by age 10, which roughly aligns with the end of the Foundation Phase (South African Government, 2019). This important yet aspirational goal is, however, set against a sobering reality. The 2016 round of the Progress in International Reading Literacy Study (PIRLS) found that 78% of South African Grade 4 readers could not reach the low international PIRLS benchmark – a signal for being able to read for meaning. This

¹ No fee schools cannot charge school fees. These are schools in quintiles 1 to 3, the system DBE used to rate schools according to the income, unemployment and literacy levels in a community. The system is used to determine public funding to schools.

² The original Early Grade Reading Study was replicated in Mpumalanga with a slightly different model. The latter is referred to as EGRS II. In this document we refer the original Early Grade Reading Study as the EGRS or EGRS I

³ In 2019 and 2020, the DBE requested USAID's support in proceeding with an expansion of the EGRS. The focus of the expansion was to scale up the coaching intervention, which showed the largest impact in the initial evaluations. The existing Reading Support Project (RSP) was therefore modified to include selected EGRS components. The RSP has since been implemented in 164 of the original 230 EGRS schools, with a further 50 schools serving as controls.

compares to 4% among grade 4 readers internationally. When broken down into language groups, almost 90% of grade 4 Setswana readers could not read for meaning (Howie et al. 2017).

To read with understanding in African languages or English, various foundational reading subskills need to be mastered before children can, when reading on their own, 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. Without suitable assessment data to measure these skills, problems acquiring them can go unnoticed, with undeveloped decoding skills showing up only later in very poor written comprehension as reflected in PIRLS results. The poor identification of decoding skills is further perpetuated through a lack of agreed standards as to what constitutes an on-track reading development trajectory across different African or home languages or among learners having to acquire second language reading proficiencies.

What are reading benchmarks and thresholds?

Reading benchmarks and thresholds, provide standards against which teachers can measure learners’ reading subskills. These numerical measures of proficiency in specific reading skills can be used to monitor whether children are on track to be able to read with fluency and understanding. They can also be used to identify early on learners who are at risk of not learning to read for meaning by age 10, highlighting where effective remediation should take place (Jukes et al. 2020).

In processes to establish reading benchmarks, it is important to decide which reading subskills should be benchmarked. Developing skills in all subcomponents of reading are important but too many benchmarks can be confusing and hard to track. This report focuses on just two: letter-sound knowledge as a basic skill which refers to alphabetic knowledge of the written code; and oral reading fluency (ORF) referring to the ability to read words in context with speed, accuracy and prosody. These three components to ORF are defined here as follows:

- Accuracy refers to the percentage of words attempted 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).

Assessed measures of ORF typically focus only on speed and accuracy because prosody is subjective and difficult to measure. In this report, the term fluency is used to describe reading with speed and accuracy

and is measured as the number of correct words per minute (abbreviated as ‘cwpm’) read from a passage of text.

The purpose of the proposed letter-sound and fluency thresholds and benchmarks is outlined below:

- **Letter-sound benchmark** - identifies whether learners are developing sufficient alphabetic knowledge that underpins decoding skills necessary for accuracy in reading.
- **Oral Reading Fluency (ORF) thresholds** - identify learners who are entering an emergent level of fluency which supports reading accuracy but which is not yet sufficient to support reading with understanding. Reading below a threshold impedes reading development.
- **ORF benchmark** - identifies a minimum fluency level that is necessary for learners to comprehend what they are reading and articulates to teachers a point at which they should concentrate on further developing comprehension skills.

Report aims

The first half of this report is dedicated to establishing letter-sound benchmarks and oral reading fluency thresholds and benchmarks for EFAL speakers in the early grades. This work complements recent research efforts to establish early grade reading benchmarks in Nguni languages (Ardington et al., 2020, 2021a) and Setswana⁴ (Wills et al., 2022) in line with aims to establish benchmarks in all official South African languages (DBE, 2020b).

African home language serves as the language of learning and teaching (LOLT) until the third year (or end of the Foundation Phase) in most South African schools. Then LOLT switches to English from Grade 4 while home language instruction continues. Whereas just a quarter of learners in South Africa are taught in English in the Foundation Phase, 90% are instructed in English from Grade 4 onwards.⁵ This report focuses on EFAL, establishing fluency thresholds and benchmarks which are contextually relevant for second language English speakers in no-fee schools in South Africa. These thresholds and benchmarks are established to track learners’ EFAL decoding skills not only in the Foundation Phase but into the

⁴ Wills et al, 2022. The Setting Benchmarks in Setswana Home Language Report, and the combined full technical report is available on the USAID DEC <https://dec.usaid.gov/dec> and the Department of Basic Education Research Repository(<https://www.education.gov.za/ResearchRepository.aspx>).

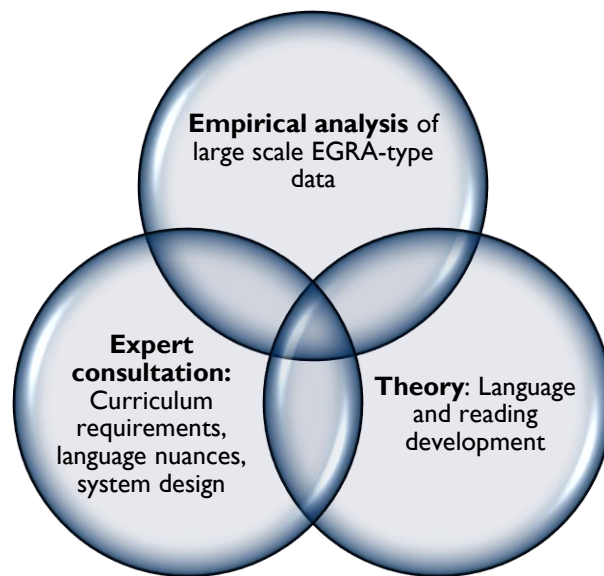
⁵ Estimates from the Annual National Assessments of 2013. See <https://www.bridge.org.za/wp-content/uploads/2016/06/SPAULL-2016-BRIDGE-reading-presentation.pdf>

Intermediate Phase, given current evidence revealing slow reading development well into the higher primary grades.

Data and approach

The context-specific reading thresholds and benchmarks established in this report result from the intersection of three main research components: A review of the theory of language and reading development, empirical analysis of large-scale early grade reading assessment (EGRA) type data and expert consultation.

Figure E 1: Three research components intersect to establish early grade reading thresholds and benchmarks



Any efforts to establish reading thresholds and benchmarks should be grounded in theory about language and how reading develops in different languages. Theory then provides the backbone for an exploratory analysis of decoding skills and their interrelationships, using large-scale early grade reading assessment data.

Key to establishing empirically driven thresholds and benchmarks, is an examination of the relationship between accuracy and speed in reading and then fluency and reading comprehension. These relationships have been well established in alphabetic languages including English (Deno et al., 2001). Accuracy in recognising letters and words has been shown to develop first, and once accuracy is established, reading rates increase as children's mastery of reading increases (Fuchs et al., 2001; Spear-Swerling, 2006). However, the nature of these relationships has been understudied in South African languages. We are sensitive not to impose assumptions about what these relationships look like in Setswana, and rather allow

an analysis of empirical regularities and reading trajectories to identify critical thresholds and benchmarks in decoding skills.

Traditional approaches to benchmarking reading subskills often focus on identifying a single point or benchmark where decoding skills are sufficiently established to support comprehension (Abadzi, 2012). However, drawing on a ‘threshold hypothesis’ by Wang et al. (2019), reaching fluency levels as defined by a benchmark may only be attainable once a minimum threshold of proficiency in fluency has developed. Identifying both thresholds (also referred to as grade specific minimum benchmarks) and benchmarks can be a more informative approach to guide reading development and reading instruction. Using longitudinal data, we establish the predictive validity of fluency thresholds for meeting fluency benchmarks (and higher levels of comprehension) in later grades.

We examine the relationship between accuracy and speed in reading and then fluency and reading comprehension, acknowledging that accuracy develops more slowly in English than in African languages (Katz & Frost, 1992). After identifying an EFAL fluency threshold and benchmark through exploratory analysis of this data, they are aligned to grade levels by investigating their attainability against current reading norms, while carefully considering grade specific reading norms for English second language speakers from other international contexts (see for example Anderson, 1999; Jimerson et al. 2013).

From research conception to final report writing, we have ensured that the established reading thresholds and benchmarks in EFAL are contextually relevant through collecting or using suitable existing EGRA-type assessment data for learners in no-fee schools. The data used are described below:

- **EFAL reading data:** The largest existing source of data on South African learners’ decoding skills in EFAL is compiled for this report by combining the EGRS I data⁶ with English assessment data from other reading studies including Story Powered Schools (SPS), Leadership for Literacy (LFL), Funda Wande in Limpopo (FW LP) and the second Early Grade Reading Study (EGRS II) (see Menendez & Ardington, 2018; Wills & van der Berg, 2020; Ardington & Henry, 2021; Department of Basic Education & University of Witwatersrand, 2020 for more information on these datasets). This culminates in a dataset with multiple EFAL assessment points for about 23,000 unique learners from Grades 2-7, across 6 of 9 provinces. These data are almost exclusively

⁶ 5 waves of EGRS I data have been collected to track learner outcomes over time. The new data for this study (wave 5 data) includes EGRS I and RSP impact evaluation and benchmarking data for EFAL and Setswana HL.

drawn from no-fee schools so that resulting benchmarks are relevant to millions of South African learners.

These data are used not only to establish benchmarks and thresholds, but to carefully examine whether they are attainable by learners at the grades for which they are set, yet aspirational enough for learners to be on a successful reading trajectory. The contextual relevance of the reading thresholds and benchmarks is also established through regular consultation with linguists, home language Setswana specialists and officials from the Department of Basic Education (DBE).

Through this research process, which brings together theoretical knowledge on reading development, empirical analysis, and expert knowledge of language and system realities, we recommend the following reading thresholds and benchmarks in EFAL.

Early grade reading thresholds and benchmarks: English First Additional Language (EFAL)

Thresholds and benchmarks for learners to develop adequate reading skills in what is typically the language of learning and teaching from Grade 4, are presented below.

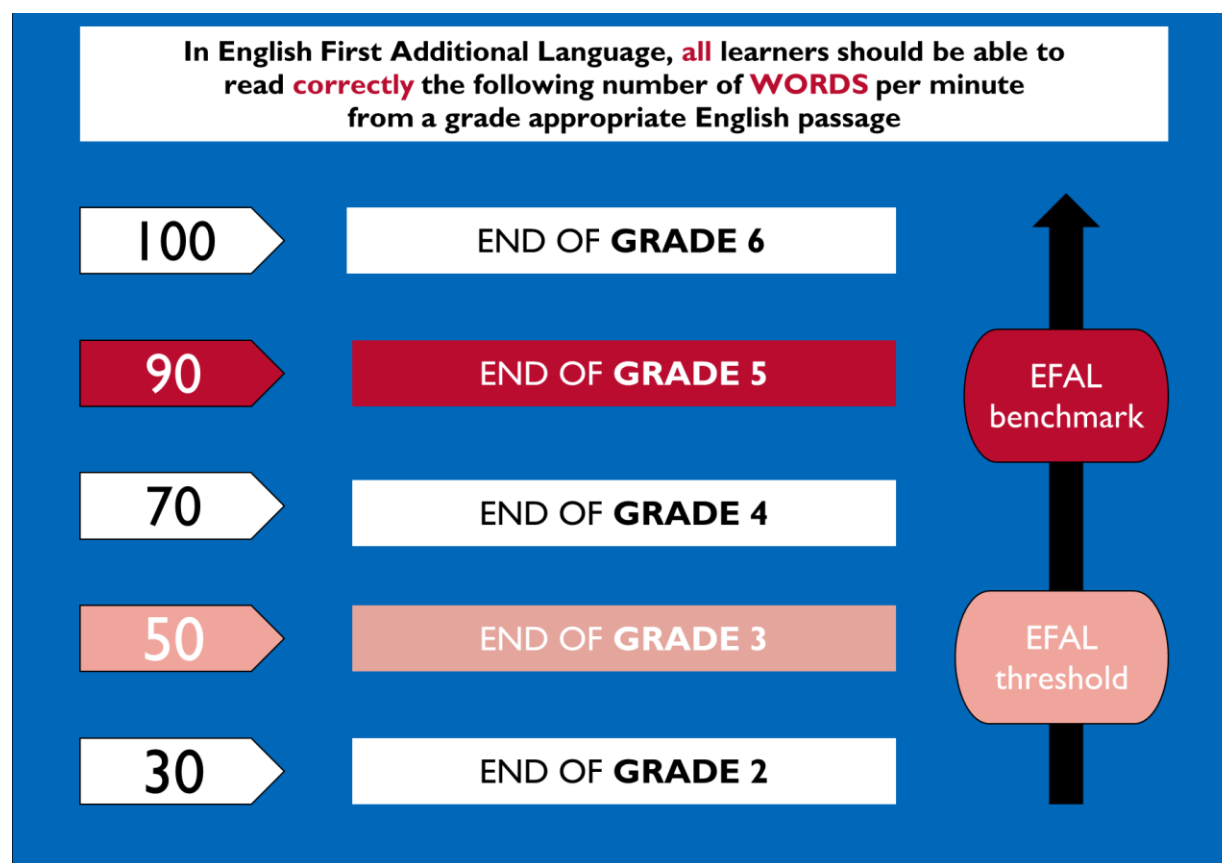
Our exploratory analysis of empirical regularities across large-scale EFAL reading data from no-fee schools reveals that below 50 words per minute, EFAL reading speeds are not only slow, but reading is inaccurate. There are very few learners reading at speeds less than 50 words per minute that read with 85% accuracy (i.e., get 85 of every 100 words correct). However, even if learners get all words correct when reading at speeds below 50 words per minute, they are reading too slowly to grasp the meaning of what they read.

Once learners start reading at speeds of around 50 words per minute, higher levels of accuracy (85% or more) are achieved. When learners are getting 50-89 words correct per minute, comprehension shifts into a development zone, where typically learners correctly answer about 4-6 of every 10 comprehension questions about a passage. By the time learners are reading 90 correct words per minute (cwpm), comprehension has reached much higher levels but comprehension improvements tend to diminish above 90 cwpm. Across various grades and passages, we typically find that only learners reading at or above 90 cwpm are getting most questions correct about the passage they have read.

From this exploratory analysis we derive an EFAL fluency threshold of 50 correct words per minute (cwpm), and a benchmark of 90 cwpm. When considered against reading norms of second language English speakers from other international contexts, and EFAL reading levels among South African learner samples

in no-fee schools, the fluency benchmark of 90 cwpm in South Africa can be reasonably aligned with the Grade 5 level. We work forwards and backwards from this point to align thresholds and benchmarks to Foundation and Intermediate Phase grades, starting with Grade 2 when learners are introduced to group guided reading and ending with Grade 6, corresponding with the last year of the Intermediate Phase.

Figure E 2: Fluency benchmarks and thresholds for early grade reading in EFAL



At the end of Grade 2, all EFAL learners should be reading grade-appropriate English passages with fluency levels at or above 30 cwpm.

- This should be viewed as an early steppingstone to meeting the EFAL fluency threshold of 50 cwpm by the end of Grade 3. Reading at 30 cwpm is far too slow to support comprehension yet it is predictive of whether learners can reach a fluency threshold of 50 cwpm a year later. In our large data set we find very few learners reading less than 30 cwpm reaching the fluency threshold of 50 cwpm a year later.
- By the end of Grade 2, median fluency among learners in Mpumalanga is just 11 cwpm. However, among those who can read one word correctly, median fluency is 20 cwpm. Reading at 30 cwpm is not entirely out of reach by the end of Grade 2 if decoding skills are effectively taught.

Table E 1: EFAL fluency levels at the end of each primary school grade among no-fee learner samples from 6 provinces (EGRA-type studies from 2017-2021)

Grade level	% unable to read one word correctly	Median fluency if non-readers are included in samples	Median fluency if non-readers are excluded from samples	EFAL fluency level to be met by <u>all</u> learners
2	21	11	20	30+
3	19-34	13-34	27-45	50+
4	8-30	21-46	44-50	70+
5	6	62	65	90+
6	1-8	80-97	84-97	100+
7	9	89	95	

Sources: Own calculations from EGRS I, RSP, EGRS II, SPS, LFL and FW-LP. Notes: Reading progressions by grade are varied and not as uniform as expected due to varied text difficulty, variations in protocols used to administer assessments and due to pandemic related impacts on reading development with some assessments conducted in 2020 and 2021.

By the end of Grade 3, all EFAL learners should be reading grade-appropriate English passages with fluency levels at or above the EFAL threshold of 50 cwpm.

- When learners reach this *minimum* threshold of 50 cwpm, they start to derive a limited amount of meaning from what they are reading. Below this Grade 3 fluency threshold, learners cannot comprehend what they have read, revealed in very low oral or written comprehension scores.
- Meeting this fluency threshold of 50 cwpm is strongly associated with whether learners will be reading at the EFAL benchmark of 90 cwpm in later primary grades. However, meeting this fluency threshold is a necessary but not sufficient condition for meeting the benchmark. Decoding skills must be taught, and learners should practice reading daily to ensure that fluency develops further.
- By the end of Grade 3, between 17% and 37% of non-repeating learners across various samples had reached this Grade 3 threshold.
- Yet far too many children (19-33% across samples) are currently not even able to read one word correctly in English by the end of Grade 3.
- Moving towards all Grade 3 second language English speakers reading at least at the fluency threshold of 50 cwpm, will require a refocus on teaching and practicing decoding skills in the Foundation Phase and ensuring that every child meets the Grade 2 fluency development threshold of 30 cwpm.

By the end of Grade 4, all EFAL learners should be reaching 70 cwpm when reading grade-appropriate English passages.

- Although international English Second Language (ESL) norms suggest that learners should be reading at the benchmark of 90 cwpm by the end of Grade 3 or 4 (Anderson, 1999; Jimerson et al. 2013), realistically this is currently not within reach for all South African learners. Among learner samples from no-fee schools in Mpumalanga, the Eastern Cape (EC), KwaZulu-Natal (KZN) and North West (NW) province, just 8-15% meet the benchmark of 90 cwpm by the end of Grade 4.
- We therefore propose a lower development benchmark of 70 cwpm for the end of Grade 4. This a steppingstone to reaching the EFAL benchmark of 90 cwpm by the end of Grade 5.
- Achieving a target where all learners are reading at least 70 cwpm by the end of Grade 4 will require significant increases in fluency in earlier grades. Across samples, median fluency at the end of Grade 4 ranges from 21-46 cwpm, or 44-50 cwpm if non-readers⁷ are excluded from samples.

By the end of Grade 5, all EFAL learners should be reading grade-appropriate passages at or above the EFAL benchmark of 90 cwpm.

- Reading at this fluency level is necessary to be able to reach higher levels of comprehension.
- However, fluency levels at or above 90 cwpm are necessary but not sufficient to support improvements in comprehension.
- This milestone signals the point at which learners become receptive to the teaching of the skills and strategies they need to tackle written comprehension and teachers should continue to encourage vocabulary and language development.
- By the end of Grade 5, 29% of a rural no-fee learner sample in KZN or the EC meet the benchmark of 90 cwpm and median fluency is 62 cwpm.

By the end of Grade 6, all EFAL learners should be reaching 100 cwpm when reading grade-appropriate English passages.

- In international studies, ESL reading around 100 cwpm is typically reported at the 50th percentile in Grades 5-6. For home language English speakers, however, fluency at the 50th percentile among sixth-grade students is around 146 cwpm in the United States (Hasbrouck & Tindal, 2017).

⁷ Unable to read one word correctly.

- By the end of the Intermediate Phase (Grade 6), across South African learner samples from no-fee schools, 38-56% are meeting the EFAL benchmark of 90 cwpm. Median fluency ranges from 80-97 cwpm so that 100 cwpm is attainable if fluency levels are strengthened in earlier grades in line with grade-specific fluency thresholds and benchmarks.

Although a Grade 7 minimum fluency level has not been provided, fluency and comprehension skills should also continue to be taught and encouraged into the final year of primary school as reading development for many learners is very slow. For example, by the end of primary school (Grade 7), 37% of the EGRS 1 learner sample in North West province do not meet the Grade 5 EFAL benchmark of 90 cwpm. About 11% are reading below the EFAL fluency threshold of 50 cwpm, indicating that they are reading too slowly to be in a zone where they can understand much at all.

While these grade specific, minimum EFAL fluency levels may be criticised as being too low by international standards, it is important to qualify that we are not presenting these as desired levels of learner performance at the 50th or 75th percentile. Rather they **reflect minimum fluency levels to be attained by all learners in no-fee schools by the end of each grade**. As reading improves into the future, these minimum suggested EFAL fluency levels should be shifted upwards. In schools or classes where EFAL reading development occurs faster, teachers and parents can refer to available international English fluency norms as guides. For example, in the United States, at the Grade 6 level, 141-159 cwpm is a fluency range at which learners are at minimal risk of reading failure and the corresponding range at the Grade 3 level is 114-135 cwpm (University of Oregon, 2021:123-124).

We also emphasise that reaching these contextually relevant minimum EFAL fluency levels should not be viewed in isolation of reaching home language reading thresholds and benchmarks. The development of decoding skills in African home language reading provides an important foundation for learning to read in English as both African languages and English are alphabetic languages. The Curriculum Assessment Policy Statement (CAPS)⁸ for EFAL develops from the assumption that when children begin to read and write in their additional language, they already know how to decode in their home language. It assumes that they have grasped concepts of print and have prior knowledge of sound-spelling relationships (DBE, 2011).

⁸ CAPS is a comprehensive policy for grade and subject specific learning, teaching, recording assessments and reporting in South African schools.

Recommendations

This research significantly advances efforts to establish reading benchmarks and thresholds in English First Additional Language (and African languages). The body of research will need to be extended until reading skills are benchmarked in all official South African languages. This will require ongoing reading assessment initiatives to expand the available set of data to establish new language benchmarks while validating and testing existing benchmarks with different learner samples and different reading passages.

In contributing to furthering benchmarking initiatives, all new early grade reading data collection initiatives should be preceded by rigorous instrument design and piloting processes to ensure appropriate EGRA-type instruments are administered. In this regard, this EFAL benchmarking exercise has established best practices to follow through highly collaborative and iterative processes of instrument development. Furthermore, following from the Nguni languages early grade reading benchmark report (Ardington et al., 2020, 2021a), the exploratory statistical methods applied to establish thresholds and benchmarks along a reading development cline are shown here to be applicable to English second language and Sesotho-Setswana reading.

Moving forward, the value of early grade reading benchmarks in supporting improvements in policy and practice will be further realised once they are linked to a national programme to assess and monitor early grade reading skills (Ardington et al., 2021a:14). Early grade reading assessments (EGRA) should form a critical part of formative assessments in primary schools. The progressive roll-out of EGRA training for teachers by the Department of Basic Education in 2015 should be leveraged to promote EGRA testing in all schools (Maboya, 2020). As teachers conduct EGRA-assessments, guided by threshold and benchmarks, this will help them to identify early-on whether learners are on track, and align their instructional practice with each learner's level of reading development.

Furthermore, if national EGRA-type assessments were introduced, for example through linking this to Systemic Evaluation Programme plans at the Grade 3 and 6 level (DBE, 2020a), thresholds and benchmarks could be used to monitor sector progress in reading in the early grades. Yet merely measuring reading skills will be insufficient for progressive improvements. Addressing the very slow reading or non-development of decoding skills among learners requires significant action and changes in practice at various levels.

Ensuring that all teachers are equipped with the knowledge and resources to effectively teach and assess reading in home languages and English is critical. Both in-service and pre-service training programmes that promote best practice in reading instruction should continue to be tried, tested and evaluated. The

growing problem of large class sizes in the early grades also needs to be addressed as individualised reading instruction or even assessment is hindered when class sizes exceed prescribed recommendations (DBE, 2020a:106). More effort also needs to be given to considering how reading can receive higher priority in homes so that children are more exposed to oral language and print at earlier ages. Children are entering school with underdeveloped emergent literacy and language skills (Dawes et al., 2017), making the work of teachers and acquiring decoding skills much harder. There is also little evidence to suggest that decoding skills are being introduced effectively in Grade R.

Finally, resources urgently need to be allocated so that reading remediation programmes are available in all schools, and that schools are equipped with enough personnel and reading resources to support not just a few learners, but many. This is particularly necessary in a context where COVID-19 schooling disruptions have impacted severely on learner's reading development, particularly in the early grades (Ardington et al., 2021b).

I INTRODUCTION

This Comprehensive Technical Report recommends English First Additional Language (EFAL) reading benchmarks for adoption by the South African Department of Basic Education (DBE). The report presents pertinent literature that informed the benchmarking study, it describes the benchmarking methodology in detail, and details an analysis of data collected from 225 no-fee schools⁹ in North West schools, between September and November 2021.

In the 2016 round of PIRLS Literacy assessments, 78% of South African Grade 4 readers could not read for meaning, as opposed to 4% of Grade 4 readers internationally. When broken down into language groups, almost 90% of Grade 4 Setswana readers could not read for meaning (Howie et al. 2017). When such large numbers of children cannot understand what they read after four years of schooling, even though most children (70%) were assessed in their home language¹⁰, action needs to be taken to understand why children cannot read and to identify what can be done to support improved reading development.

One contributing reason for poor reading comprehension in both home language and first additional language is that children have not learnt foundational skills required to be able to read and understand what they are reading. Knowledge is required of the ‘code’ of the language in which learners are reading, which we refer to as decoding skills. Reading comprehension assessments such as the Progress in International Reading and Literacy Study (PIRLS) identify higher order reading skills that many learners have not mastered. However, written language assessments such as PIRLS 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. These assessment data can be used not only to assess levels of decoding skills, but they can also be used to set contextually appropriate milestones to track children’s reading development. We refer to these milestones as reading thresholds and benchmarks.

⁹ No fee schools cannot charge school fees. These are schools in quintiles 1 to 3, the system DBE used to rate schools according to the income, unemployment and literacy levels in a community. The system is used to determine public funding to schools.

¹⁰The mismatch between being assessed in a language different from that spoken at home was largest for the English readers. In the African languages this was much lower (Howie et al. 2017).

It is not sufficient, however, to merely establish standards to support reading development in home language. African home language (HL) language of teaching and learning (LOLT) is maintained until the third year (or end of the Foundation Phase) in most South African schools. Then the language of teaching and learning switches to English from Grade 4 and mother tongue is taught as a subject rather than used as the LOLT. Whereas just a quarter of learners in South Africa are taught in English in the Foundation Phase, 90% are instructed in English from Grade 4 onwards.¹¹ It is therefore vital that on-track reading development occurs in both home language and EFAL in the Foundation Phase. If children cannot read for meaning in home language and English by the end of Grade 3, their ability to access the curriculum (or read textbooks) is significantly impeded. Yet little guidance has been provided to South African teachers around reading benchmarks to support reading development in EFAL, despite existing evidence and theory on reading in English.

Box 1: What are reading benchmarks and why are language specific benchmarks needed?

Reading benchmarks and thresholds are numerical measures of proficiency in specific reading skills that can be used to monitor whether children are on track with their reading. Benchmarks and thresholds can inform a **shared vision** of what successful reading looks like at specific grades. They provide a **standard** against which teachers can measure learner's 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. They serve as a form of quality control within an education system so that large numbers of learners do not fall through the cracks (Jukes et al. 2020).

Benchmarks for reading exist in many other languages and countries, with well-defined oral fluency norms and benchmarks for early grade reading in English home language (Hasbrouck & Tindal, 2006, 2017). Due to differences in the phonological, morphological, and orthographical features of African languages, one cannot simply apply English reading benchmarks to other languages. Furthermore, across African language groups, one cannot apply the benchmarks from one language group such as Nguni languages to others such as Sotho languages. At the most basic level, it makes no sense to impose the same fluency standards across languages with vastly different word lengths (Spaull, Pretorius & Mohohlwane, 2020). Beyond that, one needs to allow for language specific accuracy-speed and fluency-comprehension relationships that reflect reading development. Despite advances in establishing benchmarks in Nguni languages, until now, detailed work on the development of reading benchmarks in Sotho languages has not been done. Furthermore, there are no existing fluency standards in the South African context for EFAL.

Source: Setting Reading Benchmarks In South Africa https://pdf.usaid.gov/pdf_docs/PA00X1NZ.pdf.

1.1 BACKGROUND

Despite the government of South Africa's (GoSA) large investment in basic education, the country continues to face challenges providing a quality education in the majority of schools and its education indicators continue to lag behind those of its peers. In international comparative reading tests, South Africa

¹¹ Estimates from the Annual National Assessments of 2013. See <https://www.bridge.org.za/wp-content/uploads/2016/06/SPAULL-2016-BRIDGE-reading-presentation.pdf>

consistently performs at the bottom with nearly 80 percent of Grade 4 learners unable to read with comprehension in the language of their choice including home language (Howie et al, 2016). The GoSA considers education to be one of its highest domestic priorities and one of the greatest long-term challenges facing the country, as is evident in the National Development Plan which states its number one objective as improving the quality of basic education (DBE, 2013).

To support the GoSA, USAID Southern Africa, awarded the PERFORMANCE Indefinite Delivery Indefinite Quantity (IDIQ) to Khulisa Management Services (Khulisa) to provide technical, analytical, advisory, monitoring, evaluation and related support services to assist USAID Southern Africa in **effectively diagnosing needs, and planning, designing, monitoring, evaluating and learning from interventions**. PERFORMANCE helps to fill a critical research gap by providing rigorous analysis in target areas related to improving the quality of language and literacy skills of primary grade learners in South Africa and the region. USAID Southern Africa found there is little data available on the impact of teacher training programs on student learning outcomes, including literacy. Additionally, there is little known about fluency and reading benchmarks for learning in African languages. PERFORMANCE aims to fill this crucial research gap with rigorous evaluations, studies and assessments. Task Order 4 under PERFORMANCE has 12 objectives, three of which relate to establishing learning benchmarks:

- Objective 6: Collect Setswana and EFAL benchmarking data in Grade 6 (and Grades 5 and 4 as may be the case in the existing sample for EGRS), leveraging the EGRS impact data collection effort. Top up the sample of learners if necessary to meet the sample size requirements.
- Objective 8: Collect Setswana and EFAL benchmarking data in Grade 3 (and Grades 2 and 1 as may be the case in the existing sample for RSP) and in Grade 2 (from the new sample), leveraging the RSP impact data collection effort. Top up the sample of learners if necessary to meet the sample size requirements set out below in the section “Evaluation design and methodology”
- Objective 11: Conduct data cleaning and analysis of all relevant data collected in Grades 1-6 to establish reading benchmarks in Setswana and EFAL, in close collaboration with the DBE research team. Provide a high-quality technical report and learning brief in which benchmarks are proposed to the DBE for adoption.

To address the challenge of children not learning to read for understanding, in 2015 the DBE initiated the first Early Grade Reading Study (EGRS I)¹² in two districts in the North West province (districts of Ngaka Modiri Molema and Dr Kenneth Kaunda). The EGRS evaluated three Setswana Home Language (HL) interventions aimed at improving reading in the early grades: a teacher training intervention, an on-site teacher training and coaching intervention, as well as a parental intervention. The interventions showed significant impacts on learner results when teachers benefitted from training, coaching, and provision of learning materials. In 2018, data was collected from the same sample of learners. This EGRS I evaluation showed that the initial impacts of the EGRS on learners' ability to read, continued one year beyond the end of the intervention.

In 2019 and 2020, the DBE requested USAID's support in proceeding with an expansion of the EGRS I. The focus of the expansion was to scale up the coaching intervention, which showed the largest impact in the initial evaluations. The existing Reading Support Project (RSP) was therefore modified to include selected EGRS components. The RSP has since been implemented in 164 of the original 230 EGRS schools, with a further 50 schools serving as controls.

The purpose of Task Order 4 is to contribute to the body of research around early grade reading interventions, to make progress towards establishing reading benchmarks in South African Home Languages and English First Additional Language

1.1.1 Aims

In this report, we use existing adapted early grade reading assessment (EGRA) type data and new data collected through the EGRS I and RSP in North West Province to establish letter-sound benchmarks and oral reading fluency thresholds and benchmarks for the early grades in home language Setswana. In addition, this data is combined with English assessment data from multiple reading studies (including Story Powered Schools, Leadership for Literacy, Funda Wandé and the second Early Grade Reading Study) across five other provinces to establish reading thresholds and benchmarks in English First Additional Language (EFAL) for second language English speakers. These data are all drawn from no-fee (Quintiles 1-3) schools so that resulting benchmarks are relevant to the majority of South African learners who are attending less resourced school contexts. This research serves to complement an emerging body of

¹² The original Early Grade Reading Study was replicated in Mpumalanga with a slightly different model. The latter is referred to as EGRS II. In this document we refer the original Early Grade Reading Study as the EGRS or EGRS I.

research to establish early grade benchmarks in all official languages in South Africa (see Jukes et al., 2020; Ardington et al., 2020, 2021a).

For the purposes of this benchmarking exercise, we define fluency as the ability to read with speed and accuracy. Although prosody is a component of fluency, it is difficult to measure consistently and reliably in field studies and thus is not considered here. The purpose of the proposed letter-sound and fluency benchmarks is outlined below:

- Letter-sound benchmark - identifies whether learners are developing sufficient alphabetic knowledge that underpins foundational decoding skills necessary for accuracy in reading.
- Fluency threshold - identifies learners who are entering an emergent level of fluency which supports reading accuracy but which is not yet sufficient to read for meaning.
- Oral Reading Fluency (ORF) benchmark - identifies a lower bound fluency level that is necessary for learners to comprehend what they are reading and articulates to teachers a point at which they should concentrate on further developing comprehension skills.

Aligned to three points in the Foundation Phase, we establish a Grade 1 letter-sound benchmark, a Grade 2 fluency threshold and a Grade 3 fluency benchmark for home language Setswana speakers. For English second language speakers we establish a Grade 3 fluency threshold and Grade 5 fluency benchmark. We also work forwards and backwards from these points to establish other grade-specific minimum fluency levels in the Foundation and Intermediate Phases that act as steppingstones or further developmental goals for EFAL reading fluency in primary school. These thresholds/benchmarks should be viewed as establishing the standard or level that every learner should attain at grade specific points to move through a successful language and reading trajectory journey in the primary school years.

These reading thresholds and benchmarks are not determined in an arbitrary manner. Scientific literature on reading across language groups should inform benchmarks. They should be based on strong empirical work and should be sensitive to current realities of learning and curriculum requirements. Our data approach to establish benchmarks in this study is guided by language and reading theory, as well as expert advice. The established threshold and benchmarks are necessary to get learners onto a successful reading trajectory without being so aspirational that no-one can reach them.

Following Ardington et al. (2020), our approach to setting thresholds and benchmarks integrates theoretical understandings of reading development with a non-parametric analysis of cross-sectional and

longitudinal EGRA-type data. The proposed minimum grade specific EFAL fluency levels, are also examined against existing norms for English second language (ESL) speakers in other international contexts.

1.1.2 Report structure

Before detailing our empirical method and results, the next section (section 2) turns to reading development theory that guides our empirical analysis. We provide a theoretical framework for reading development, reviewing three models of early reading development to see how they account for successful reading and the processes that drive it. We also propose an overarching developmental framework of early reading against which early reading across all the official languages in South Africa can be mapped with language specific developmental differences identified. Given that reading theories pertain to both Setswana and English.

EFAL (Section 4) commences with a discussion of evidence on early reading development in ESL or EFAL. The methodology applied in EFAL benchmarking process, the data used and how the benchmarks are aligned to grade levels is presented in Section 6. The results showing how an EFAL threshold and benchmark are developed are discussed Section 7, then Section 8 outlines the rationale for how they are aligned to Foundation and Intermediate Phase grade levels.

The summary of EFAL thresholds and benchmarks discussed in Section 9. The concluding section considers how to increase the efficacy of established reading thresholds and benchmarks through shaping improved policy and practice for improved reading outcomes.

2 THEORETICAL FRAMEWORK FOR EARLY READING DEVELOPMENT

The goal of reading is to understand or comprehend what we read. Yet children in South Africa are struggling to develop this key skill in both home language and English – a skill that is necessary for learners to be able to access the curriculum. To understand why our children are struggling to become readers we need to understand how children become successful readers and what kinds of knowledge and competencies they require to develop on the journey of being able to read with meaning.

Reading itself is a complex process. It requires the development of various skills and knowledge factors including knowledge of linguistic factors¹³, text factors¹⁴, code-based factors¹⁵ and the development of a range of lower to higher level cognitive processes. The components needed to become a skilled reader do not necessarily develop simultaneously. Some aspects of the phenomenon may develop first, laying a foundation on which subsequent competencies are built and thrive (Stanovich 2000). It is estimated to take several hundred hours of practice over the years, through regular exposure to written language, to become a skilled reader. Additionally, the extent to which reading development occurs is likely to be supported or impeded by external factors such as socio-economic factors, the home environment, the schooling context, instructional practice, access to books, nutrition, health, affection, and emotional security.

In the past 70 years, scientific theories of reading have been strongly influenced by reading research centred on English and other Western European languages.¹⁶ Yet in the past thirty years, a broader body of evidence available from different linguistic and writing systems is starting to inform current theories of reading. In particular, findings from studies on reading and particularly early reading in *transparent* orthographies and in other *agglutinating languages* (e.g., Finnish, Turkish, Basque) may be of relevance to early reading in African languages since they share common orthographic and linguistic features.

2.1 READING THEORY: THREE RELEVANT MODELS

2.1.1 Simple view of reading

The simple view of reading (SVR) posits that in order to understand a text, children need to develop in two main skills areas, namely decoding skills (knowledge of the written code) and oral language proficiency (knowledge of the language in which they read) (Gough & Tunmer 1986, Hoover & Gough, 1990). Ability in both skills is necessary for reading comprehension but each are not on their own sufficient to be able to read with understanding. Both skills are important, but in the initial stages of reading development, decoding is a critical skill that is necessary for children to read a text on their own and make sense of it. Once children have acquired relative mastery over the code and can read words accurately and fluently,

¹³ Knowledge of the language of the text at a sub lexical (phonological and morphological), lexical (vocabulary), sentence (morphology and syntax) and discourse-level.

¹⁴ Knowledge of text conventions; the functions that different genres of text serve; the way information is structured in different genres within and across paragraphs; the role of headings; visuals, etc. Text or topic complexity, topic familiarity and word frequency levels can also affect reading comprehension.

¹⁵ The technical features of a specific writing system and the linguistic features of the language onto which the written symbols are mapped.

¹⁶ Including German, Dutch (Germanic language family), French, Spanish, Italian (Romance language family)

then its influence diminishes and other skill areas related to language proficiency drive reading comprehension, such as vocabulary.

Basically, the SVR predicts that early skilled readers (e.g., in Grade 1) will show strong decoding skills as well as strong language proficiency. Struggling readers in Grade 1 will not have adequate decoding and/or adequate language proficiency to enable them to read and to understand a text on their own. In Grade 1 there will be variations in decoding skills between readers but by the end of Grade 3, when learners are expected to read fluently on their own with understanding, such differences should have more or less levelled off as mastery in decoding is established. From Grade 3 onwards variations in language proficiency (e.g., vocabulary differences) and other cognitive processes (such as making inferences during reading) should account for variations in reading comprehension.

Although there is converging evidence across alphabetic languages of the need for strong decoding skills to enable comprehension, in line with the simple view of reading, there are nuanced differences as once decoding is well established, differences in language proficiency (e.g., in vocabulary or morphosyntactic knowledge) become stronger predictors of reading comprehension. For example, in Turkish and Finnish (both transparent, agglutinating languages), accuracy in letter-sounds and word reading is achieved early (end of Grade 1) and remains high (Durgunoğlu & Öney 1999; Leppänen et al. 2008), while in English this develops more slowly. In transparent codes, accuracy is reached relatively early and so decoding as measured by word accuracy is not as strongly related to reading comprehension as linguistic comprehension. However, when decoding is measured by fluency, then its importance in reading comprehension is found to extend to later grades (Florit & Cain 2011), although these patterns can be influenced through pedagogical approaches in the classroom. For example, because accuracy is achieved so early in Finnish, reading rate (fluency) thereafter distinguishes weaker readers from normally developing ones.¹⁷

2.1.2 The decoding threshold hypothesis

The SVR assumes a linear relationship between decoding and reading comprehension, i.e., the stronger a child's decoding skills are, the more likely they will understand what they read. More recently, Wang, Sabatini, O'Reilly, and Weeks (2019) have proposed a refinement of the SVR by arguing that the

¹⁷ Reading in Finnish, in all schools, is taught via a systematic phonics approach, all the letters are taught in the first term of Grade 1, teachers are well qualified, all early grade teachers are trained in the phonics approach and teach it consistently, schools are well resourced, and struggling readers are identified and remediated early (Aro 2017). See also Kendeou et al. 2013.

relationship is more complex. They posit a decoding threshold hypothesis, which predicts that there are ‘threshold conditions’ in the relationship between decoding and reading comprehension. In other words, 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). They base their claims on the analysis of very large longitudinal data sets in English (over 30,000 learners). 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.

The findings from Ardington et al. (2021a) based on a large dataset of 14,000 readers in three Nguni languages found evidence that learners reading below a decoding threshold were in a non-comprehension zone. They also find that beyond a certain point, increases in fluency do not translate into higher comprehension.

2.1.3 Orthographic depth hypotheses

According to the orthographic depth hypothesis (ODH), the reading process will develop slightly differently depending on the orthography of the language (Katz & Frost, 1992). Alphabetic writing systems can be placed on a continuum of transparency in the way in which sounds are mapped onto letters. Languages with transparent systems have a fairly regular mapping between sounds and letters (i.e., the same letter symbol always represents the same sound) whereas languages with more opaque writing systems have more irregular letter-sound mappings. For example, African languages have transparent (or shallow) orthographies, whereas English has a more opaque (or deep) orthography. Early reading success in alphabetic orthographies is expected irrespective of whether they are transparent or opaque (Alcock et al. 2010; Melby-Lervåg, Lyster, & Hulme, 2012). However, the ODH predicts a quicker process of learning to read in transparent languages than opaque orthographies. Due to regular mapping, learning to read in a transparent orthography will happen more quickly, and novice readers will rely more on phonological processing (i.e., direct phoneme to grapheme mappings).¹⁸

¹⁸ For example, children learning to read in transparent orthographies seem to do so more quickly than children learning to read in English (Seymour et al. 2003). The ODH predicts a different trajectory for children learning to read in an opaque orthography: Due to irregular mapping systems, novice readers will take longer to learn to read and will rely not only on phoneme-to-grapheme mapping but will also utilise a lexical strategy (like whole word mapping) to read words.

Another extension of the ODH theory, called the Psycholinguistic Grain Size Theory (PGST) predicts that reading in transparent orthographies will be easier by comparing ‘grain size’ across languages. ‘Grain size’ refers to units that readers use to decode words, which can vary from whole word, syllable, morpheme or phoneme-to-grapheme mappings. The PGST proposes that readers of transparent orthographies rely on the mainly small grain size of letter-sounds to read words while readers of opaque orthographies rely also on larger grain sizes such as rimes (e.g., -at as in *cat*, *mat*, *sat*), syllables or even whole words to read words because of irregularities in the phoneme-grapheme mapping system.

Scientific evidence also seems to support that reading development occurs faster in transparent languages.¹⁹ Based on findings from other transparent languages and agglutinating languages, accuracy in alphabetic knowledge tends to occur early and mastery level (accuracy and fluency) in both letter-sound knowledge and word reading should be achieved within the foundational years of schooling. As discussed later, however, this advantage might be offset by the more complex consonant sounds that occur in African languages. Cross-linguistic details still need to be further investigated.

2.1.4 Instructional practice shapes reading development

So far, we have focused on three theories of reading. It is also important to acknowledge that early reading development across language types is likely to be strongly influenced by instructional practice, although studies seldom investigate the role that instruction plays on early reading trajectories.

Early reading instruction in English has been beset by the reading wars, where whole language approaches downplay the role of code-based knowledge and phonological processing, while the phonics approach stresses the importance of explicitly teaching children how the code works and how to blend letter-sounds in sequence so that they can decode words. Such differences in instructional context may subtly affect the early reading trajectories of English readers, for example, novice readers in whole language classrooms may exhibit stronger lexical strategies for decoding because phonological processing is downplayed. It is important, therefore, to examine early reading trajectories in light of possible influences of instructional practices.

¹⁹ Studies have shown that in transparent languages such as Greek, Finnish and Turkish, children master the code by the end of Grade 1 and rely strongly on phonological processing to do so.

2.2 A DEVELOPMENTAL VIEW OF MULTIPLE PROFICIENCIES

Reading development is dynamic and changes over time such that readers' profiles look different at different points in development in transparent and opaque orthographies. Different code- and cognitive-based processes dominate or diminish in importance and are superseded by others as proficiency increases. A skilled Grade 1 reader displays a different profile from a skilled Grade 4 reader (Adams 1990; Stanovich 2000; Castles et al. 2018; Kim 2020; Caravolas et al. 2021).

What this basically means is that a skilled Grade 1 reader is somewhat different from a skilled Grade 4 reader; what manifests as reading competence in the early years of schooling changes from what manifests as reading competence a few years down the line.

Yet, there is converging evidence from different languages which points to decoding accuracy developing first (e.g., in letter-sound knowledge, syllable and word reading), followed by increased processing speed, which then leads to automaticity in processing (i.e., processing without effort or conscious attention). This frees up working memory and attention for meaning construction. The arrow in Figure 1 depicts the direction of this developmental trajectory. However, the points or thresholds at which accuracy or increased alphabetic knowledge lead to automaticity in word reading (in or out of context), thereby enabling reading comprehension, may differ across languages depending on their linguistic and orthographic features.

Figure 1: Developmental cline in early reading



The influence of some processes as drivers of reading development change as proficiency increases and their role is taken over by qualitatively different processes. Drawing on the work of Stern et al. (2018), different types of readers are distinguished on a cline, specifically because there are no clear-cut differences between them. This is described more fully in Box 1. The development framework emphasises that 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. The three studies on early reading in Setswana discussed in section 4 show that this is not yet happening in Grade 2 or 3. This might help to understand why the PIRLS outcomes in Setswana are so low.

In the report that follows, we aim to show in greater detail what code-based thresholds support reading development and reading comprehension in Setswana as a home language followed by English as a second language.

Different types of readers can be distinguished on a continuum.

Non-readers show poor print awareness, poor phonological awareness¹ and have very little 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.

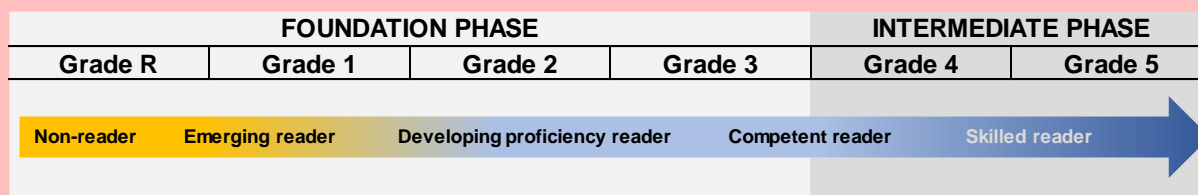
Readers in the next phase are **developing proficiency**. They have more accurate and fluent knowledge of the alphabetic code, which enables them to decode syllables and words in or out of context with greater accuracy (approximating 95% accuracy in home languages which refers to correctly reading 95 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.

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.

The figure below loosely maps the reading skill cline against grade progression in the early school years.

Developmental continuum in early reading in relation to formal grades



By the time learners exit 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 *non-readers* or *emerging readers* at the end of grade 3. Such a developmental lag would point to challenges in pedagogy and the teaching/learning context.

3 BENCHMARKING METHODOLOGY

3.1 APPROACH

As explained in the theoretical overview, 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. The importance of some earlier processes diminish as proficiency increases and they are replaced by qualitatively different processes. Yet across all the components, accuracy develops first, followed by increased processing speed, which then leads to automaticity (processing without effort or conscious attention) as discussed in section 2.2. This automaticity free-ups cognitive resources (e.g., working memory and attention) to be allocated to constructing meaning from text (LaBerge & Samuels, 1974; Samuels & Flor, 1997, Fuchs *et al.*, 2001; Spear-Swerling, 2006). Following this developmental cline, our empirical work will first examine the relationship between accuracy and speed and then the relationship between fluency and comprehension.

Our approach draws on the idea of non-linearities in the relationship between decoding and reading comprehension where comprehension only occurs above a certain level of decoding proficiency (Wang *et al.* 2019, Kim 2017, Kim & Wagner 2014). In other words, if decoding skills are below a minimum threshold, reading comprehension remains stagnant. As mentioned above, Wang *et al.* (2019) speculate that there may also be an upper threshold, beyond which there are no additional gains to comprehension from increased decoding skills. Our empirical work will focus on identifying these critical threshold points in learners' reading development.

A significant body of evidence demonstrates that reading processes differ by language, with variations in the core skills employed by children learning to read (Katz & Frost, 1992; Torppa *et al.*, 2016; Dowd & Bartlett, 2019). This calls for language-specific benchmarks. It also implies that we should be cautious not to impose assumptions about the specific nature of the speed-accuracy or fluency-comprehension relationships in Setswana in our empirical approach.

Our benchmarking approach is also mindful of how benchmarks and thresholds need to be contextually valid. They should not be set so high as to be out of reach for the majority of early grade learners. At the same time, they need to be ambitious enough to establish expectations that are sufficient to support meaningful improvements in early grade reading. They also need to be appropriate for the curriculum context and aligned to the priorities of teaching across school system phases. Our approach to setting benchmarks is anchored to the context through data, with statistical methods that also support contextually relevant thresholds/benchmarks as explained in the next section. Additionally, we have also

assessed whether identified thresholds and benchmarks are contextually relevant through consultation with linguistic and curricula experts, including home language Setswana speaking literacy experts.

3.2 STATISTICAL METHOD AND SKILLS

3.2.1 Exploratory non-parametric techniques

Following Ardington et al. (2020, 2021a) and Jukes et al. (2020), we engage in exploratory data analysis, guided by reading development theory and expert linguistic knowledge. To avoid imposing *a priori* assumptions about the nature of understudied reading development in Setswana, we use non-parametric techniques to explore the accuracy-speed and fluency-comprehension relationships with the purpose of identifying critical points in learners' reading trajectories.

Once potential thresholds are identified, we test them to establish whether these critical points provide meaningful distinctions between learners and whether they align with the stages of reading development. We do this using concurrent data (data from the same grade-point) on related reading skills. We also investigate whether the proposed thresholds are set at levels that can be achieved by current learners; sensitive to incremental changes in reading performance in this context and, at the same time, ambitious enough to support meaningful improvements in reading proficiency. In other words, both backwards and forwards analysis of the data are used in a system of checks and balances to verify the results.

3.2.2 Advantages of non-parametric methods over other benchmarking approaches

Widely-used approaches to benchmarking rely on identifying a fixed comprehension threshold (e.g., at least 80% of questions correct) and then apply statistical techniques to identify the fluency levels associated with meeting that comprehension threshold (Room to Read, 2018; Abadzi, 2012; RTI, 2010). Our approach to benchmarking has some advantages over these previously used methodologies (see Ardington et al., 2021a; Jukes et al. 2020).

First, non-parametric methods make no assumptions about the speed-accuracy or fluency-comprehension relationships which can be affected by the linguistic differences between languages and pedagogy. A pedagogy that focuses little on teaching comprehension skills can also result in lower comprehension scores in that context which are independent of students' reading skills.

Second, our thresholds are invariant to the serious challenges of establishing the appropriate level of comprehension questions. Traditional benchmarking methods assume that a fixed level of comprehension is a comparable construct across passages and languages. Our established thresholds are invariant to the challenges of cross-text comparability of comprehension questions or the impact of the placement of the

comprehension questions in relation to the text. We illustrate the challenge of establishing comprehension difficulty using the Setswana data in the analysis that follows.

Third, our approach does not depend on having large samples of learners with advanced comprehension skills (Abadzi, 2012). Research on Nguni languages shows the relationship between fluency and comprehension flattens out at fairly low levels of comprehension. This suggests that beyond a certain fluency point, poor comprehension skills become the limiting factor and we may find insufficient numbers of learners achieving proficient (e.g., 80%) comprehension levels to support the identification of benchmarks. Our approach is not dependent on a fixed level of comprehension, so that low comprehension skills are not a limitation. Our identification of critical thresholds in the accuracy-speed and fluency-comprehension gradients relies on examining the full distribution of these relationships whereas traditional methods only focus on these relationships around the specific comprehension cut-off.

3.2.3 Establishing ORF benchmarks

3.2.3.1 *Examining the relationship between speed and accuracy*

Existing research on Nguni languages in the early grades is indicative of a non-linear relationship between reading accuracy and reading speed. Accuracy and speed initially increase together steeply but eventually accuracy does not improve as much with additional increments in speed – the relationship flattens out. The point at which this relationship tapers off can inform a reading threshold. Then examining the reading speed at which sufficient accuracy is achieved can inform what benchmark should be established.

Preliminary analysis of the wave 4 EGRS I ORF data, and pilot data, suggested a similar non-linear relationship in Setswana, however given the shorter nature of word length in the disjunctive orthography of Setswana (compared to the Nguni languages), tentative analysis suggests that the point at which 95% accuracy will be reached will be much higher than in the Nguni languages (Spaull et al., 2020). We examine this further using a larger dataset.

3.2.3.2 *Examining the relationship between fluency and comprehension*

In the same way that we examine the relationship between speed and accuracy, we then explore non-parametrically the relationship between fluency (a measure of both speed and accuracy) and comprehension. We seek to establish if there are regular patterns that exist in this relationship, and whether critical thresholds can be identified. For this analysis, we use samples that are allowed 3-minutes to read a passage and complete a subset of the ORF comprehension questions. Although we examine a fluency-comprehension relationship, our approach avoids benchmarking fluency to a fixed comprehension level.

3.2.3.3 *Validity checks – concurrent and predictive validity*

We then establish the concurrent validity of the proposed fluency threshold and benchmark by examining how they align against the performance of the same learners on written comprehension assessments. We then use longitudinal data to examine whether meeting specified thresholds or benchmarks at earlier grade points is predictive of learners' future fluency and comprehension proficiency levels. Given the low levels of reading proficiency, we use data from later grades to understand what a successful trajectory could look like, working forwards and backwards to establish the grade thresholds/benchmarks for an 'on track' successful reading journey.

3.2.4 Establishing letter sound benchmarks

To set a letter-sound benchmark in Nguni languages, Ardington et al. (2020) relied on a combination of empirical insights from a large reading dataset in Nguni languages and expert opinion of letter-sound knowledge required for mastery in decoding, grounded in the theory of reading. The authors motivate that by the end of Grade 1, learners should be reading 40 correct letter sounds per minute (clspm) or more. Like Nguni languages, Sesotho-Setswana languages are alphabetic languages. Despite differences in pronunciation, one wouldn't expect significant differences in the process of letter sound acquisition across these language groups. To find empirical support for this hypothesis, we repeat the analysis in the Nguni-language benchmarking report using Setswana data. We initially explore the relationship between accuracy and speed in sounding out letters. Then longitudinal data (data for the same learner for two timepoints) is used to examine incremental improvements in letter-sound knowledge against baseline scores to identify whether there is a point beyond which gains in letter-sound knowledge are negligible.

3.3 INSTRUMENT DEVELOPMENT PROCESS: INSTRUMENTS THAT SUPPORT BENCHMARKING

As discussed above, the choice of statistical methods and approaches used in benchmarking can shape the outcomes of that process. However, equally important is the availability of suitable assessment data with which to apply these methods.

Prior to the collection of EGRS I wave 5 data, the waves 1-4 ORF assessments were not suitable for examining the relationship between reading fluency and comprehension. To effectively model this relationship, there must be a large enough sample that read far enough in the text to be able to answer all reading comprehension questions. This requires that enough time is provided to learners to read the text, and that the text is not too long. In earlier EGRS I waves (and many other reading studies) ORF

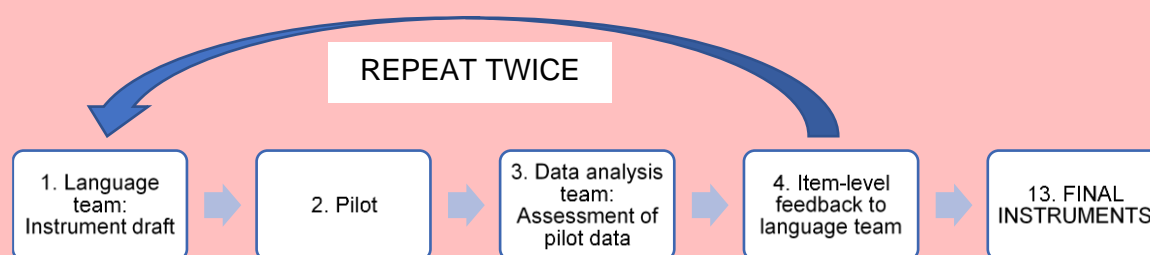
assessments only allowed reading for a minute.²⁰ As a result too few learners had the chance to answer all comprehension questions.

The instruments²¹ were piloted for data collection in 2021, which allowed for both 1-minute and 3-minute timings for the oral reading assessment. As a desired outcome of the instrument development process (and piloting iterations), most learners should be able to finish the passage within 3-minutes so they can attempt all the comprehension questions. At each stage of the piloting process (described in Box 2), the analysis team assessed passage length against the percentage of learners reading the entire passage in 3-minutes. Typically, the length of passages was found to be too long.

Box 2: Piloting process – key to establishing appropriate language assessments to support benchmarking

To improve the validity and relevance of literacy assessments used for benchmarking processes, in 2021, Khulisa Management Services conducted three pilots to determine the appropriateness of instruments developed to assess literacy skills in no-fee schools in South Africa (see the Appendix Table A 1). After each round of piloting, the data from the assessments was analysed at the item-level and presented to a language team including linguists, DBE officials and home language Setswana experts in order to guide the iterative improvement of language instruments (or assessments). The process of development involved 13 steps as shown below.

Instrument development process



The key aims of the piloting were to make sure that 1) the instruments are set at the right level with no floor or ceiling effects, 2) the language is appropriate for the context with comprehension questions asked that are unambiguous and 3) that the length of the assessment (and specific reading passages) are appropriate for the learners and the evaluation context. In particular, key purposes of the pilots were to assess the suitability of the length of the passages of connected text used to measure oral reading fluency (ORF) and to assess the difficulty level of comprehension questions. Importantly, assessments were also designed and adjusted to account for poorer reading performance expected during a pandemic period.

²⁰ Learners are then only asked comprehension questions related to the parts of the passage that they have read within the time limit. This creates an artificial relationship between fluency and comprehension.

²¹ Refer to the Report on the Development of Learner Assessment Tools and Contextual Tools (Bisgard et al., 2021)

Consequently, they were reduced with each pilot iteration and more appropriate passage lengths were finally achieved to support the availability of larger sample sizes reading the whole passage at final data collection. The piloting process was also used to assess and adjust the difficulty level of the comprehension questions asked about the passage. Pilot data was checked for irregularities in results, to highlight any instrument wording that may be ambiguous. The pilot data was also checked to ensure that the scores across all the questions aligned with expectations in terms of the comprehension processes that each question was tapping into. Various alterations were made at each piloting stage to improve the appropriateness of the comprehension questions.

In summary, three rounds of piloting integrated with detailed analytical processing of data was vital to the finalisation of the assessment instruments and the quality of the data used for benchmarking purposes. The piloting process also demonstrated the efficacy of data analysts working together with linguistics and language experts for the development of appropriate literacy assessments.

BENCHMARKING READING SUBSKILLS IN ENGLISH FIRST ADDITIONAL LANGUAGE

In Grade 4, many children start using their additional language, English, as the LOLT. As specified in the CAPS, this means “they must reach a high level of competence in English by the end of Grade 3, and they need to be able to read and write well in English. For these reasons, their progress in literacy must be accelerated in Grades 2 and 3” (DBE, 2011:8). Investigating reading assessment data for around 23,000 second language English speakers from over 900 no-fee schools, however, reveals the challenges that learners experience in developing fluency and comprehension skills in English, not only in the Foundation Phase but also into the Intermediate Phase.

In this section, we establish fluency thresholds and benchmarks in EFAL to guide reading development in the Foundation and Intermediate Phase. Before delving into an empirical analysis, we situate our approach within reading development theory and evidence of what we already know about reading in a second language, particularly from Southern Africa.

4 EARLY READING DEVELOPMENT IN ENGLISH SECOND LANGUAGE (ESL) OR FIRST ADDITIONAL LANGUAGE (EFAL)

Much of the reading research undertaken in the past 50 years has been on English, primarily on reading among English HL speakers but also increasingly on reading in English by non-HL speakers or ESL readers. The simple view of reading is strongly supported by English reading data, both when word reading or ORF is assessed. Robust relationships, ranging from moderate to strong correlations, are typically found between all the foundational reading skills, namely phonemic awareness, letter-sound knowledge, word reading, nonword reading and ORF (cf. for example Kilpatrick 2012; Castles et al. 2018). Similar trends are found in ESL reading (cf. for example Geva & Sadeh 2006; Lipka & Siegel 2007). Because this is well-trodden ground in English reading research, these issues will not be revisited in this review. Furthermore, some of these issues have already been covered in section 2.1.1 in the review of reading in Setswana, where support for the simple view of reading was briefly discussed, and in sections 2.1.3 where similarities and differences between reading in an opaque language such as English and transparent languages were briefly identified.

Instead, we focus here on identifying what research on early reading in English in a more local African context has found. However, first we turn briefly to international evidence on normative reading performance in HL English which is an important frame of reference for ESL or EFAL reading performance.

4.1 INTERNATIONAL EVIDENCE ON NORMS AND BENCHMARKS IN ENGLISH READING

HL English reading is not directly comparable to ESL/EFAL reading as HL readers always have the advantage of greater language proficiency and a larger vocabulary. However, some studies have shown that despite this advantage, HL and ESL readers can perform at similar decoding levels if systematic phonics instruction is provided (Lesaux & Segal 2003; Geva & Zadeh 2006; Lipka & Siegel 2007).

Over 25 years ago, benchmarks for letter-sound knowledge (LSK) in English were set at 40 clspm for Grade 1 (Kaminski & Good III 1996). Although this benchmark was derived from learning to read in English, it is possible that most languages that use the Roman alphabet in their orthography will reflect fairly similar LSK benchmarks (cf. Ardington et al. (2020) for a similar LSK benchmark derived independently from a dataset of about 12,000 readers in the Nguni languages).

Based on large-scale longitudinal data from over 28,000 English readers, Hasbrouck and Tindal (2006) identified ORF norms for readers from Grades 1-6. For example, by the Spring (or end) of Grade 1, readers at the 50th percentile had an ORF mean of 53 cwpm in 2006, which increased to 60 cwpm in more

recent assessments of 1-2 million children (Hasbrouck & Tindal, 2017). Using the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) framework, first language (L1) readers who read 39 cwpm at the end of Grade 1 are at minimal risk of reading failure, while this figure is 94 for the end of Grade 2 and 114 for the end of Grade 3 (University of Oregon 2021:123-124). Reading rate in English as a second language (L2) is estimated to be at about 70% of that of a first language (L1) reader (Anderson, 1999). Jimerson et al. (2013) found an average difference of around 25 cwpm between English HL and ESL readers at each grade level from Grades 1-4.

In the next section, we compare these norms against evidence of decoding skills in EFAL among students in existing Southern African studies.

4.2 EVIDENCE FROM SOUTHERN AFRICA

4.2.1 Reading levels in English

Although there is some information about foundational reading skills in English in Southern and East African schools, studies are somewhat scattered and diverse, and instruments are not standardised across the studies which makes comparisons within and across grades difficult. Fortunately, the tasks and procedures for assessments of LSK and ORF are fairly uniform across studies.²² For this reason, we reflect only on findings on LSK measured as letter sounds correct per minute (clspm) and ORF measured as correct words per minute (cwpm), drawing from Lekgoko and Winskel (2008), Piper (2009), Malda et al. (2014), Pretorius (2014), Pretorius and Spaull (2016), Zenex Literacy Project (2017), Menendez and Ardington (2018), Liswaniso (2021), Makaure (2021) and Mutema (2021). Table 1 describes in more detail the studies for which findings on ORF and LSK are summarised in Table 2.

Although the learners in these studies are classified as 'English', data from English HL learners is not always distinguished from data from learners for whom English is a second or additional language. Some of the data also include learners from schools where English was the LOLT from the Foundation Phase, while other studies have data on learners for whom the LOLT was an African HL during the Foundation Phase and who shift to English as LOLT in Grade 4.

²² Although many of the studies in this review include word reading, performance is not readily comparable across studies because the instruments used to assess word reading are very diverse, with some being untimed and some timed tests, and some including decodable and sight words in the same list, while others assess them separately. Similarly, the instruments used to assess reading comprehension are very diverse with varying length of text (e.g., written PIRLS texts), and some including only oral reading comprehension questions based on short and easy texts.

Table 1: Descriptions of Southern African studies identifying letter-sound knowledge and oral reading fluency in English among early grade readers

Author	Purpose / purpose of Study	Sample	LOLT in Foundation Phase
Lekgoko and Winskel (2008) [Botswana]	Compared the relationship of phonemic awareness and letter knowledge to word reading	35 Grade 2 children who had learned to read in Setswana (Grade 1) and then in English (Grade 2) in 2 rural state schools in Botswana. Average age 8.4 years at data collection. Assessed mid-year in Grade 2.	
Piper (2009) [South Africa]	Impact evaluation of the Systematic Method for Reading Success (SMRS) program	Baseline (start of Grade 1) and post assessment of grade 1s half-way through the year (June). Post assessments for 546 learners from about 45 schools in Limpopo, Mpumalanga and North West province.	
Malda et al. (2014) [South Africa]	The researchers were interested in comparing similarities/differences across opaque (English) and transparent orthographies (Afrikaans and Setswana)	358 Grade 3 learners in English, Afrikaans and Setswana. Tested using a battery of tests that tap into different linguistic, cognitive and code-based components of reading to identify strengths and weakness in their reading development. English results for 127 learners.	English
Pretorius (2014) [South Africa]	Reports on a catch-up reading intervention in a high poverty primary school near Pretoria with Grade 4 isiZulu HL/EFAL learners	44 learners in one school. The learners were assessed on an array of tests decoding, spelling, dictation and reading comprehension tests at the beginning and end of Grade 4 to track their reading development.	
Pretorius and Spaul (2016) – NEEDU [South Africa]	A secondary analysis of the NEEDU (National Education Evaluation and Development Unit) data	ORF and reading comprehension data were gathered from 1,772 Grade 5 EFAL learners in rural schools across various provinces.	
Piper & Zuilkowski (2015) [Kenya]	Study aims to look at critical issues of assessment, examining whether assessments should be administered orally or silently	The authors compare Grade 2 student's scores on oral and silent reading tasks of the EGRA in Kiswahili and English using data from the Primary Math and Reading Initiative in Kenya. 1541 learners randomly selected from random sample of schools in Bungoma and Machakos counties.	

The Zenex Literacy Project (Zenlit) (2017) [South Africa]	A 3-year intervention project aimed at improving the teaching and learning of reading in Foundation Phase in KwaZulu-Natal (isiZulu), the Eastern Cape (isiXhosa) and the Western Cape (English)	299 learners assessed: Grade 1 n=92, Grade 2 n= 92, Grade 3 n= 106. The English schools were low-income schools in Mitchell's Plain. Although most of the learners were LI speakers of English, they tended to speak a non-standard dialect of English characteristic of the Cape Flats. Five subcomponents of EGRA were used to assess early reading skills in the three languages.	English
Second Early Grade Reading Study (EGRS II) (waves 1-3) [South Africa]	Longitudinal evaluation focused on improving EFAL teaching and learning	A large-scale study, involving over 3,000 EFAL learners across quintile 1-3 schools in two districts in Mpumalanga Province. Learners are from 49 isiZulu and 131 Siswati schools. Data were gathered at four different time points: baseline (beginning of Grade 1), Wave 1 (end Grade 1), Wave 2 (end Grade 2) and Wave 3 (end Grade 3)	Siswati / isiZulu
Menendez & Ardington (2018) - Story Powered Schools (SPS) (2018) [South Africa]	Longitudinal evaluation of the Story Powered Schools programme	A large-scale study in KwaZulu-Natal and Eastern Cape no-fee schools. Learners assessed in isiXhosa or isiZulu using a battery of EGRA-type tasks from Grades 2-5.	isiXhosa / isiZulu
Liswaniso (2021) [Namibia]	Reading intervention aimed at improving the decoding and reading comprehension skills of the Grade 5 learners.	Grade 5 ESL learners in Namibia in 4 high poverty schools (2 control n = 149; 2 intervention n = 157) in Katimo Mulilo, over two school terms. Both ORF and reading comprehension were assessed, with pre- and post-tests administered at the beginning and end of Grade 5.	English
Makaure (2021) [South Africa]	Compared the foundation reading skills of bilingual Northern Sotho – English readers in Grade 3	Two groups of Northern Sotho bilinguals were compared: EFAL bilingual readers who had Northern Sotho as LOLT in Foundation Phase, and ESL bilingual readers who had English as LOLT in Foundation Phase. Comparisons using an array of phonological, word reading, non-word reading and ORF subtests in both languages.	Northern Sotho & English
Mutema (2021) [Zimbabwe]	Focused on reading development across Grade 3 and 4 ESL learners in four primary schools in Zimbabwe.	A written PIRLS reading comprehension test was administered to all learners (Grade 3 n = 180; Grade 4 n= 187) and ORF tests administered one-on-one to a subsample of these learners.	English

The following trends can tentatively be identified from Table 2:

- **Very low levels of alphabetic knowledge:** Although there is not much data available on LSK among English ESL/EFAL learners, the available findings show very low levels of alphabetic knowledge. This is particularly evident when compared to international benchmarks suggesting that learners should be able to sound out 40 clspm by the end of Grade 1.
- **Very low and slow reading in English in the early years:** The findings show very low and slow reading in English among ESL/EFAL learners in the early years. Again, international English ORF norms reflected in Hasbruck and Tindal (2006, 2017) provide a point of comparison. Additionally, Pretorius and Spaul (2016) provide a visual representation of how far behind Grade 5 EFAL learners in South Africa are lagging, compared to ESL readers in the United States of America (USA) as seen in Figure 2. The dotted pink line shows where average Grade 5 English HL learners perform; the staggered blue line shows where Grade 3 ESL readers perform, and the stippled blue line shows where Grade 1 ESL learners perform in the USA. South Africa's Grade 5 EFAL readers (as shown in the histogram) are reading at the same rate as Grade 1 ESL readers in the USA, with 11% of the sample of 1,722 learners unable to read a single word in Grade 5. These findings speak to poor quality teaching of reading in EFAL classrooms.

Figure 2: Distribution oral reading fluency scores (wcpm) for rural South African ESL learners relative to Broward County ESL learners (from Draper and Spaul, 2015)

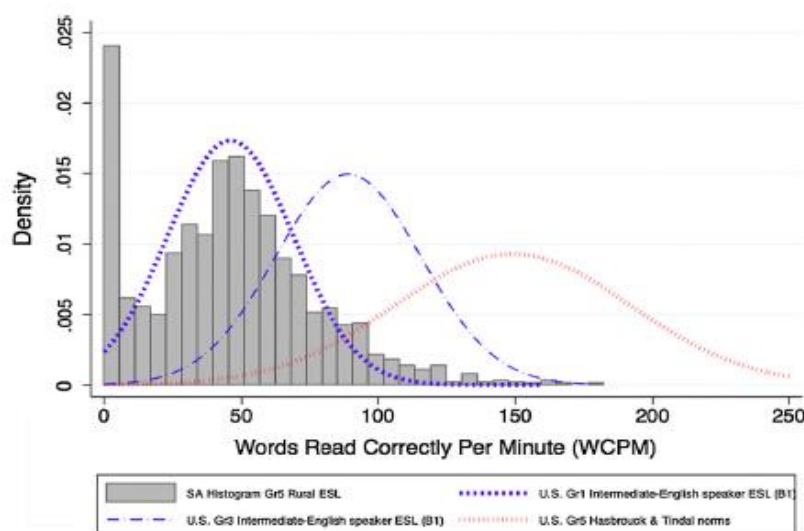


Fig. 3 Distribution of oral reading fluency scores (WCPM) for rural South African ESL learners relative to Broward County ESL learners, Florida (Broward County, 2012) (Source: Draper & Spaul, 2015)

- **EFAL learners enter Intermediate Phase with inaccurate and slow English reading ability** at a time when a double transition occurs for most learners, namely the transition from African HL to EFAL in Grade 4, and the transition from ‘learning to read’ mainly narrative texts to ‘reading to learn’ from information textbooks. Learners who have not yet acquired strong reading skills that can serve as a vehicle for learning find it exceedingly difficult to cope with these transition challenges.
- **Robust positive relationships between ORF and reading comprehension** identified across all the studies (e.g., Spearman correlations are as high as 0.70 in Pretorius and Spaul (2016); 0.73 in Liswaniso (2021); 0.78 in Grade 3 and 0.85 in Grade 4 in Mutema (2021)).²³

The study by Pretorius and Spaul (2016) using data from rural Grade 5 EFAL students collected by NEEDU suggest some tentative EFAL ORF benchmarks. Pretorius and Spaul (2016) report a levelling off in the relationship between fluency and comprehension at around 70 cwpm. They found that the small number of learners in the sample who obtained 60% or more on the reading comprehension assessment could read at least 80 cwpm but that returns to added fluency dropped thereafter. They suggest 70 cwpm as a point at which learners could start accessing meaning in texts during reading. We investigate this further using the new EGRS I/ RSP wave 5 data.

Table 2: Average letter-sound knowledge and oral reading fluency in English among early grade readers, evidence from Southern African studies

Letter Sound Knowledge	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5+
Correct letter sounds per minute (clspm)					
International norms in English Home language reading (Kaminski & Good III, 1996)	40				
Lekgogo & Winskel (2008) - Botswana		22.4			
Piper (2009)	*mid-year: 21				
Zenex (2017)	26	42.5	35.6		
Story Powered Schools (SPS)					
Baseline Eastern Cape			26	29	
Baseline KwaZulu-Natal			18	20	

²³ Although we qualify that many of these studies would have used 1 minute ORF timings and so the correlation is biased upwards.

Oral Reading Fluency	Grade 1	Grade 2	Grade 3	Grade 5	Grade 6
Correct words per minute (correct words per minute)					
International norms in English Home language reading (Hasbruck & Tindal, 2017)	60	100	112	133	146
Piper (2009)	*mid-year: 9				
Malda et al. (2014)			41.08		
Pretorius (2014)				16 →43	
Pretorius & Spaull (2016)					46.1
Piper & Zuilkowski (2015) - Kenya		30			
Zenex (2017)	17.6	44.0	55.8		
EGRS II (EFAL Mpumalanga)					
Wave 2 (int)		19.7			
Wave 2 (control)		18.9			
Wave 3 (all)			33.5		
Story Powered Schools (SPS)					
Eastern Cape				T1: 34 → T4: 43.9	
KwaZulu-Natal				T1: 39.6 → T4: 52.1	
Makaure (2021)					
EFAL (FP LOLT – African HL)			19.6		
ESL (FP LOLT – English)			24.4		
Mutema (2021) – Zimbabwe			71.3	72.9	
Liswaniso (2021) - Namibia					
Control					45.1→51.8
Intervention					41.8→57.5

Notes: T = term.

5 BENCHMARKING METHODOLOGY: EFAL

While there is little existing literature on how children learn to read in African languages, another gap in the literature pertains to learning to read in L2 in developing country multilingual education contexts. Although benchmarks for reading exist for English first language speakers (Hasbrouck & Tindal, 2006 & 2017; University of Oregon, 2021) and schemas exist for second language reading in the United States for

example (e.g., Broward County Public Schools (2012) have an ESL ORF schema), countries may take a more nuanced approach to setting benchmarks for second language reading in English. The theory of how second language reading develops, provides an important nuance to setting benchmarks in a first additional language. Although oral reading fluency in early grades in L2 is predictive of fluency in later grades (as with L1 reading), thus justifying the importance of setting second language benchmarks, the benchmarks for L2 are typically set at lower levels than for L1 (Pretorius & Spaull, 2016). Children who learn to read in L2 have less knowledge of the language, with lower oral proficiency in L2 leading to students reading text less efficiently than in their L1. Additionally how language specific features of South African languages influence how learners read in English (Katz, 2020) may have a bearing on the benchmarking process.

Following our approach to establish Setswana benchmarks (and in earlier work Nguni language benchmarks), non-parametric statistical methods are again applied to EGRA-type data to explore the relationships between reading accuracy and speed, and reading fluency and comprehension. In line with the decoding threshold hypothesis (Wang et al. 2019), we expect to observe non-linear relationships justifying the need to establish both a lower fluency threshold and upper benchmark in EFAL reading. For example, earlier work by Pretorius and Spaull (2016) on EFAL suggested a levelling-off in the relationship between fluency and comprehension at around 70 cwpm. We interrogate this further using new data.

Empirical regularities and inflection points in these relationships provide the initial basis for setting benchmarks and thresholds. However, the empirical results obtained are closely balanced against the theory of reading in a second language, curriculum requirements and other contextual realities. For example, one needs be cognisant of the very low levels of EFAL reading fluency among South African learners relative to L2 reading in other countries. We also closely weigh up our proposed fluency threshold and benchmark against evidence on L2 reading norms from other international contexts. Finally, we are conscious to set the threshold and benchmark at grade levels that are aligned to strategic curriculum grade points in consultation with representatives from the Department of Basic Education.

6 EFAL DATA

Since 2017, a growing set of English EGRA-type data has been collected for learners in largely no-fee schools. Studies that have assessed large numbers of learners in English decoding skills include the first Early Grade Reading Study (EGRS I) in North-West Province and the related Reading Support Project (RSP); the second Early Grade Reading Study (EGRS II) in Mpumalanga; the Leadership for Literacy (LFL) project in higher and lower performing no-fee schools across KwaZulu-Natal, Gauteng and Limpopo; the Story Powered Schools (SPS) impact evaluation in KwaZulu-Natal and the Eastern Cape; and the Funda

Wande (FW) evaluation in Limpopo province. More details on these studies is provided in the Appendix Table A 2. (Also see Menendez & Ardington, 2018; Wills & van der Berg, 2020; Ardington & Henry, 2021; Department of Basic Education & University of Witwaterstrand, 2020 for more information on these datasets)

The data and assessments by study, grade and term used to evaluate the relationship between reading accuracy and speed are shown in Table 3. There are assessment points for about 23,000 unique learners from over 900 schools. The data reflected is largely longitudinal, with 2 to 3 assessment points available for most learners albeit a few exceptions.²⁴

Table 3: Available samples by year, term, and grade to explore oral reading fluency in EFAL

Study	Grade	Term	Year	Text	Total words	N comprehension questions	3-minute reading	Learner sample sizes
EGRS I*	3	III	2018	a	126	7	no	2101
EGRS I	4	III	2018	a	145	8	no	3289
EGRS I	3	III	2021	b	71	10	yes	3094
RSP	4	III	2021	a	126	9	yes	3210
EGRS I	7	III	2021	c	203	9	yes	3216
EGRS II	2	IV	2018	b	70	4	no	2739
EGRS II	3	IV	2019	a	126	5	yes	2684
EGRS II	4	IV	2020	a	126	8	yes	2405
SPS Cohort II	4	I	2018	b	103	10	no	3470
SPS Cohort II	4	III	2019	b	103	10	yes	2953
SPS Cohort II	5	III	2019	b	103	10	yes	2922
LFL	3	I	2017	a	127	9	no	731
LFL	3	IV	2017	a	127	9	no	632
LFL	6	I	2017	c	203	9	no	732
LFL	6	IV	2017	c	203	9	no	599
FW LP	6	IV	2021	c	203	9	yes	1147

Notes: *This sustainability sample was added in EGRS I to assess whether teacher coaching and training impacts were sustained a year after the intervention ended. A, b, and c refer to three narrative texts (adjusted for each study), used to assess ORF across different studies. EGRS = Early Grade Reading Study; SPS = Story Powered Schools; LFL = Leadership for Literacy; FW LP = Funda Wande in Limpopo

²⁴ Exceptions include where 'top-ups' were included in EGRS I in 2021, the grade 3 EGRS I 2018 cohort which is a cross-sectional sample and where learners were not found at the next assessment period.

All the samples shown in Table 3 can be used to identify EFAL fluency levels and trajectories with ORF performance for these samples summarised in Table 8. These samples are all used to explore the relationship between reading accuracy and speed. However, the relationship between fluency and comprehension can only be explored where assessments allowed 3-minutes for passage reading (shown in green in Table 3) so that learners have enough opportunity to be able to answer all the related comprehension questions. Furthermore, the analysis of fluency-comprehension is restricted to samples where learners who read at least one word have attempted all or at least a subset of the questions as shown in Table 4.

Table 4: ORF comprehension performance for learners with ORF > 0 and allowed to read the ORF passage for 3-minutes

Study	Grade	Term	Observations with ORF>0	Learners attempting all questions				Learners attempting subset of questions			
				% of learners with ORF > 0	25th percentile comprehension	Median comprehension	75th percentile comprehension	% of learners with ORF > 0	25th percentile comprehension	Median comprehension	75th percentile comprehension
EGRS I	3	III	2117	68%	18%	36%	64%	75%	25%	38%	63%
RSP	4	III	2229	60%	22%	44%	67%	80%	14%	43%	71%
EGRS I	7	III	2914	80%	44%	67%	89%	80%	44%	67%	89%
EGRS II	3	IV	1986	71%	18%	36%	64%	71%	25%	38%	63%
EGRS II	4	IV	1831	69%	25%	38%	63%	72%	14%	43%	71%
SPS	4	III	1444	84%	10%	20%	30%	84%	10%	20%	30%
SPS	5	III	1460	91%	10%	20%	50%	91%	10%	20%	50%
FW LP	6	IV	1060	75%	44%	56%	78%	75%	44%	56%	78%

Notes: EGRS = Early Grade Reading Study; SPS = Story Powered Schools; LFL = Leadership for Literacy; FW LP = Funda Wande in Limpopo

6.1 EFAL ORF PASSAGE DIFFICULTY

English oral reading fluency data for the Grades 2-7 learner samples is drawn from three narrative texts used across these studies, which differ by word length (ranging from 70-203 words) and the number of ORF comprehension questions asked (4-10). We briefly examine the difficulty level of these English texts by subjecting different passages to three different readability tests as discussed in Box 3. When replacing any names used to refer to people in the passages with a familiar word, all the texts according to the Dale-Chall test (1948) are identified as 4th Grade or below (i.e., they are very easy to read). Text b, administered to Grade 2 and 3 learners, is identified as being set at a Flesh-Kincaid Grade Level 2. Text a and c are assessed as being at a Flesh-Kincaid Grade Level of about 2.5 to 3.3. The Dale-Chall test (1948) is arguably a better indication of readability as it accounts for word familiarity rather than just syllables, word counts and sentences.

Box 3: Three different readability tests

Flesch–Kincaid grade level. Reflects what educational level a person will need to be able to understand a particular text by applying a formula that weights total words, total sentences, and total syllables.

The Flesch Reading Ease Score. This score reflects the approximate educational level a person will need to be able to read a particular text easily. How comprehensible a document is will be indicated on the Flesch Reading Ease Score by a number between 0 and 100. Scores around 100 mean the document is extremely easy to read, while scores around 0 mean that it is highly complex and difficult to understand. In the United States, a score between 90-100 is associated with a 5th grade level or being easily understood by an average 11-year-old student.

Dale-Chall readability score: This test measures a text against a list of words that are familiar to fourth-grade students and younger in addition to accounting for total words, syllables and sentence counts. It is arguably the most suited for assessing texts used in ORF assessments.

Table 5: Readability of select EFAL passages

	Jabu and the dog	How the elephant got its trunk	Unbelievable night
	Text b: Administered to Grade 2s & 3s ^a	Text a: Administered to Grade 4 learners ^b	Text c: Administered to Grade 6-7 learners ^c
Flesch-Kincaid Grade Level	1.7	3.3	2.5
Flesch Reading Ease Score	97.3	87.7	89.8
Flesch Ease Score: Equivalent reading Level	5th Grade (Very easy to read)	6th Grade (Easy to read)	6th Grade (Easy to read)
Dale-Chall Readability Score (replace names in text with "cat") ^d	4th Grade and below (very easy to read)	4th Grade and below (very easy to read)	4th Grade and below (very easy to read)
Average Words per Sentence	7.9	9	7
Average Syllables per Word	1.2	1.3	1.3
Sentences	9	14	29
Words	71	126	203

Notes: ^aUsed in EGRS I for Grade 3, 2021; used in EGRS I for Grade, 2 2018. ^bUsed for Grade 4 RSP, 2021 & Grade 4 EGRS II, 2020. ^cUsed for Grade 6 learners in LFL (2017) and FW Limpopo (2021) and Grade 7 EGRS I (2021) learners. ^dThe following website was used to calculate the Dale-Chall Readability Score: <https://www.textcompare.org/readability/new-dale-chall-score/>

6.2 EFAL SCHOOL SAMPLE CHARACTERISTICS

The EFAL learner samples are located across 944 schools in six provinces, with six different African home languages represented across the sample of largely second language English speakers. This is implied in Table 6 which shows the characteristics of the learners' schools by study. The schools are almost all no-fee (Quintile 1-3) and a large proportion of the schools (56-100%) are in rural areas. The LOLT in these schools in the Foundation Phase is almost always an African language with Sesotho-Setswana and Nguni language teaching (and by implication Nguni and Sesotho-Setswana home language speakers) well represented across the studies. In the case of the LFL Study, English is the LOLT for 18% of the school sample but this does not necessarily imply that all learners are English home language speakers.

Table 6: School characteristics of the EFAL data sample

		SPS	EGRSII	LFL	EGRSI*	FW LP
Quintile	Quintile 1	54%	54%	23%	48%	13%
	Quintile 2	33%	33%	36%	29%	52%
	Quintile 3	12%	13%	31%	23%	35%
Location	Rural	100%	92%	56%	77%	86%
Province	EC	48%				
	GP			34%		
	KZN	52%		34%		
	LP			31%		100%
	MP		100%			
	NW				100%	
LOLT	isiXhosa	48%				
	isiZulu	52%	27%	54%		
	Siswati		73%			
	Xitsonga			16%		
	Sepedi			11%		100%
	English			18%		
	Setswana		1%		100%	
	Schools	357	180	61	230	116

Notes: *The RSP school sample is a subset of these EGRS I schools. EGRS = Early Grade Reading Study; SPS = Story Powered Schools; LFL = Leadership for Literacy; FW LP = Funda Wande in Limpopo

6.3 EFAL READING NORMS IN NO-FEE SCHOOLS IN SIX PROVINCES: ORAL READING FLUENCY AND COMPREHENSION

Table 8 provides a summary of EFAL ORF performance for the learner samples reflected in Table 3. In addition to this table, the percentage scoring zero on EFAL ORF is illustrated in Figure 3, but sorted by study so that cross-grade performance in the same schools or for the same learners over time can be tracked. From these data, we observe the following trends:

- **There are significant percentages of EFAL non-readers at the end of the Foundation Phase, and reading is very slow:** At the end of Grade 3, one fifth to a third (19-33%) of learners cannot read one word in English. Among those that can read a word, median fluency ranges from 27-45 cwpm. The prolific problem of non-reading emerges in earlier grades, and should be detected and remediated. At the Grade 2 level, observed from the EGRS II study in Mpumalanga, 21% are unable

to read one word.²⁵ Among those Grade 2s that can read, ORF is around 20 correct words per minute (cwpm).

- **The problem of non-reading in English persists into the first two years of the Intermediate Phase, despite a switch to English as the language of learning and teaching.** In the EGRS I/ RSP and EGRS II schools, about 23-30% of Grade 4s in term 3 or 4 are unable to read one word. Among those that can read at least one word, median ORF ranges from 44-50 cwpm. The SPS sample of learners in KZN and EC tend to do better. At the start of Grade 4, 11% cannot read one word with ORF typically at 40 cwpm among those reading at least one word. Over a year, the percentage of non-readers only declines to 8% and reading 49 cwpm is typical of those who can read at least one word. By the end of Grade 5, among this longitudinal sample, 6% cannot read one word correctly and the median ORF is 65 cwpm among those reading at least one word.
- **The development of EFAL reading fluency continues into Grade 6 and 7, although there remain far too many learners reading too slowly by the time they are about to transition into secondary school.** At the end of the Intermediate Phase (Grade 6), 1-8% of learners from the LFL and FW Limpopo samples cannot read one word, and median ORF among those reading at least one word ranges of 81-97 cwpm. The EGRS I Grade 7 sample, having experienced two disruptive pandemic years, share a similar reading profile to the Grade 6 samples with median ORF of 95 cwpm for those reading at least one word. About one out of every ten (9%) learners in this sample are about to enter secondary school completely unable to read at all in English, yet requiring instruction in Grade 1 or 2 level EFAL decoding skills. This should not be viewed as a COVID related problem. This sample should have been reading in EFAL by the end of the Foundation Phase, in years before the pandemic.
- **With slow English reading in the Foundation Phase and Grade 4, comprehension skills have not developed for most learners.** Across the full Grades 4 samples, at the median learners are getting 0-13% of the ORF comprehension questions correct, with percentages correct at the 75th percentile ranging from 20-50%. Restricting samples to those reading at least one word, at the median the percentage of questions correct ranges from 10-25%. Although comprehension performance is highly sensitive to the difficulty of the passage and related comprehension questions, a unanimous

²⁵ 12% of EGRS II grade 2 learners have an ORF of 1. It is not clear why this data anomaly arises, and may account for the deterioration in non-reading (i.e., ORF = 0) between grade 2 and 3 for these samples.

finding is that ORF comprehension levels are highly underdeveloped. **Consistent with PIRLS, we find that most learners cannot read for meaning in EFAL by the end of Grade 4.**

- **As English reading fluency develops into Grade 6 and 7, comprehension skills appear to follow yet remain weak.** Among the Grade 6 and 7 samples assessed at the end of the year, at the median the percentage of ORF questions learners get correct ranges from 38-44%, with scores at the 75th percentile of between 50-78%.

Although reading development in EFAL is very slow, and far slower than what is required to keep pace with the South African schooling curriculum, we also acknowledge the significant challenges that limited instructional time for English in the Foundation Phase presents for learners. In Grades 1 to 2, learners are allocated 2-3 hours per week for EFAL (see Table 7). In Grade 3, this increases to 3-4 hours per week. One potential reason for the notable improvements in EFAL fluency observed in the Intermediate Phase is that more instructional time per week (5 hours) is allocated to English. Reading development in EFAL is also more likely to flourish as home language decoding skills become better established by the Intermediate Phase.

Table 7: Weekly time allocation to First Additional Language vs. Home Language in CAPS

	Home language (HL)	First Additional Language (FAL)	Mathematics	HL time as % of total instructional time	FAL time as % of total instructional time
Grade 1	8 (7)	3 (2)	7	35% (30%)	13% (8%)
Grade 2	8 (7)	3 (2)	7	35% (30%)	13% (8%)
Grade 3	8 (7)	4 (3)	7	32% (28%)	16% (12%)
Intermediate Phase (Grade 4-6)	6	5	6	22%	18%
Senior Phase (Grade 7)	5	4	4.5	18%	15%

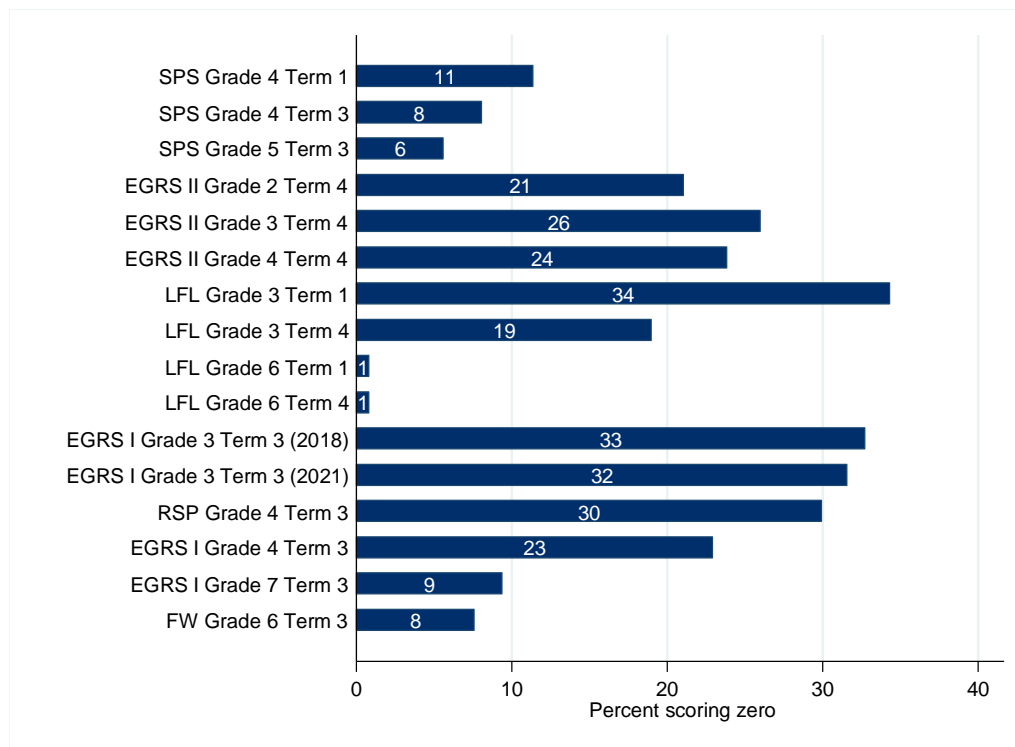
Source: DBE, 2011. Notes: Total instructional time per week is 23 hours in Grade 1-2, 25 hours in Grade 3, and 27.5 hours in Grades 4-7.

Table 8: ORF and ORF comprehension scores, South African early grade reading assessments (2017-2021)

Grade	Study	Year	Term	Assessment	All learners								Learners with ORF > 0							
					ORF				ORF comprehension				ORF				ORF comprehension			
					% scoring zero	25th percentile	Median	75th percentile	% scoring zero	25th percentile	Median	75th percentile	25th percentile	Median	75th percentile	% scoring zero	25th percentile	Median	75th percentile	
2	EGRS II	2018	IV	I	21%	1	11	32	39%	0%	25%	50%	6	20	37	23%	25%	50%	50%	
3	LFL	2017	I	I	34%	0	15	38	59%	0%	0%	11%	15	27	51	38%	0%	11%	22%	
3	EGRS I	2018	III	I	32%	0	13	38	68%	0%	0%	14%	13	27	50	53%	0%	0%	14%	
3	EGRS I	2021	III	I	32%	0	14.5	42	14%	0%	9%	36%	13	32	52	13%	9%	27%	55%	
3	LFL	2017	IV	2	19%	11.5	34	62	38%	0%	11%	33%	24	45	70.5	24%	11%	11%	33%	
3	EGRS II	2019	IV	2	26%	0	19	49	34%	0%	0%	40%	13.5	34	58	34%	0%	20%	60%	
4	SPS Cohort II	2018	I	2	11%	15	36	54	30%	0%	10%	20%	23	40	57	25%	10%	10%	20%	
4	EGRS I	2018	III	I	23%	1	30	66	51%	0%	0%	25%	22	48	74	36%	0%	13%	38%	
4	RSP	2021	III	3	30%	0	21	57	18%	0%	11%	44%	18	44	68	18%	11%	22%	56%	
4	SPS Cohort II	2019	III	I	8%	25	46	69	31%	0%	10%	20%	31	49	72	24%	10%	10%	20%	
4	EGRS II	2020	IV	2	24%	1	30	64	24%	0%	13%	50%	19	50	71	23%	13%	25%	50%	
5	SPS Cohort II	2019	III	3	6%	39	62	92	23%	10%	20%	40%	43	65	93	18%	10%	20%	40%	
6	LFL	2017	I	I	1%	59	80	107	9%	13%	25%	50%	59	81	108	8%	13%	25%	50%	
6	LFL	2017	IV	2	1%	68	97	121	5%	13%	38%	50%	68	97	121	4%	13%	38%	50%	
6	FW LP	2021	IV	I	8%	44	80	105	11%	22%	44%	67%	58	84	107	4%	33%	56%	78%	
7	EGRS I	2021	III	3	9%	57	89	119	2%	22%	44%	78%	67	95	122	2%	33%	56%	78%	

Notes: *This sustainability sample was added in EGRS I to assess whether teacher coaching and training impacts were sustained a year after the intervention ended. EGRS = Early Grade Reading Study; SPS = Story Powered Schools; LFL = Leadership for Literacy; FW = Funda Wande in Limpopo

Figure 3: Percentage of learners scoring zero on ORF assessments in EFAL, South African early grade reading assessments (2017-2021)



Notes: COVID affected samples include EGRS II Grade 4 (term 4), EGRS I /RSP Grade 3 (term 3 sample), EGRS I Grade 7 (term 3), RSP Grade 4 (term 3) and the FW Limpopo Grade 6 (term 3) sample.

7 BENCHMARKING RESULTS: EFAL

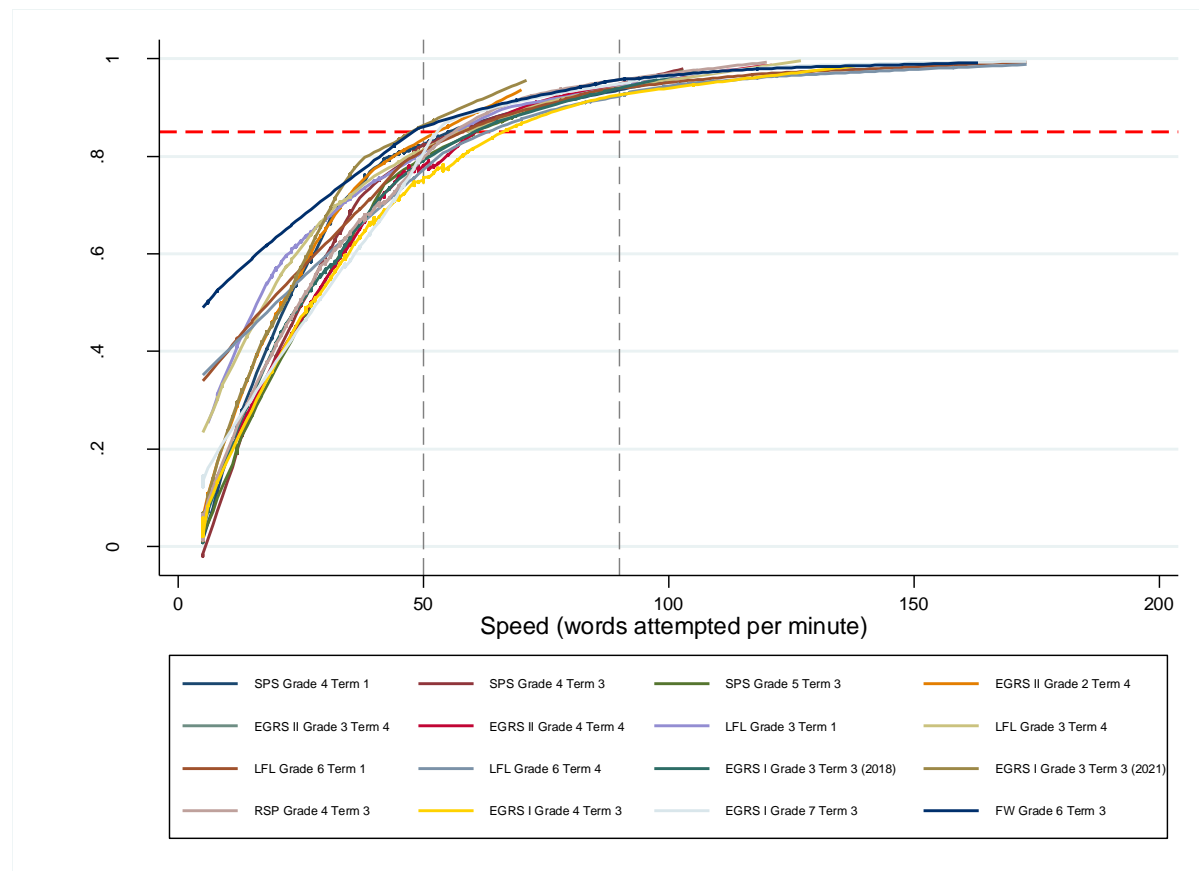
7.1 ESTABLISHING A FLUENCY THRESHOLD AND BENCHMARK IN EFAL: AN ANALYSIS OF READING SPEED AND ACCURACY

Following the structure of the exploratory analysis for Setswana reading, we initially examine the relationship between reading speed and accuracy in EFAL. Once again, reading speed is measured by the number of words attempted (in the time limit) from a passage. Accuracy is measured by the percentage of words attempted that learners' get correct. A departure from the home language benchmarking analyses for Setswana and Nguni languages – which are transparent languages - is that we use 85% as a signal for developed accuracy in the early grades in EFAL reading, rather than 95%. The reason is that children learning to read in transparent orthographies do so more quickly and achieve greater accuracy

more quickly than children learning to read in English (Seymour et al. 2003).²⁶ Furthermore, accuracy is likely to be achieved later in a second language.

Figure 4 plots the average accuracy associated with each level of reading speed using local polynomial weighted regressions. The figures include grey dashed vertical reference lines at 50 and 90 words per minute. A red horizontal line representing a minimum desired level of accuracy is set at 85% (i.e., for every 100 words attempted, the learner gets 85 correct).

Figure 4: Reading speed and accuracy in EFAL, by sample



²⁶ Studies have shown that in transparent languages such as Greek, Finnish and Turkish, children can master the code by the end of Grade 1 and rely strongly on phonological processing to do so. In their study, accuracy in English took longer, well into Grade 2. In addition to phonological processing, English readers also develop lexical processing routes to recognise words (facilitated by the fact that an analytic language like English has many short words at high frequency levels that don't really change their shape (except for verb or plural endings) - but lexical routes will only develop if children are actually getting regular opportunities to read.

A remarkably consistent concave pattern again emerges across all samples, grades and reading passages: initially accuracy and speed increase steeply together and then the relationship tends to flatten off when accuracy levels reach around 85%, and then flattens even further when accuracy levels of 95% are reached. A speed of roughly 50 words per minute aligns with the point at which this relationship starts to flatten at around 85% accuracy. When 95% accuracy is reached, learners across all samples are reading at a speed of about 90 words per minute.

We then identify the distribution of speed for accurate readers and compare it against the speed of inaccurate learners using box plots. The lower and upper line of the box indicate the 25th and 75th percentile – i.e., 50% of each learner sample have English reading speeds in this band. The median is indicated by the line in each box. We consider what these distributions look like where accurate readers are distinguished from non-accurate readers using an 85% accuracy cut-off points.

In almost all cases, the 25th percentile of reading speed for accurate readers lies at or above the line representing a reading speed of 50 words per minute (see Figure 5). A similar result is seen if we use a cut-off of 95% accuracy. In the case of inaccurate readers (less than 85% accuracy), the majority are reading with speeds less than 50 words attempted per minute (see Figure 6) (albeit the LFL samples being an exception) with all almost all learners in every sample reading substantially slower than 90 words attempted per minute.

This supports the idea that learners who read less than 50 words per minute have not yet reached accuracy levels to support automaticity. They would benefit from instruction focused on improving their decoding skills and fluency.

Contract Number:, Order Number:

Figure 5: Speed distribution for learners reading with at least 85% accuracy in EFAL

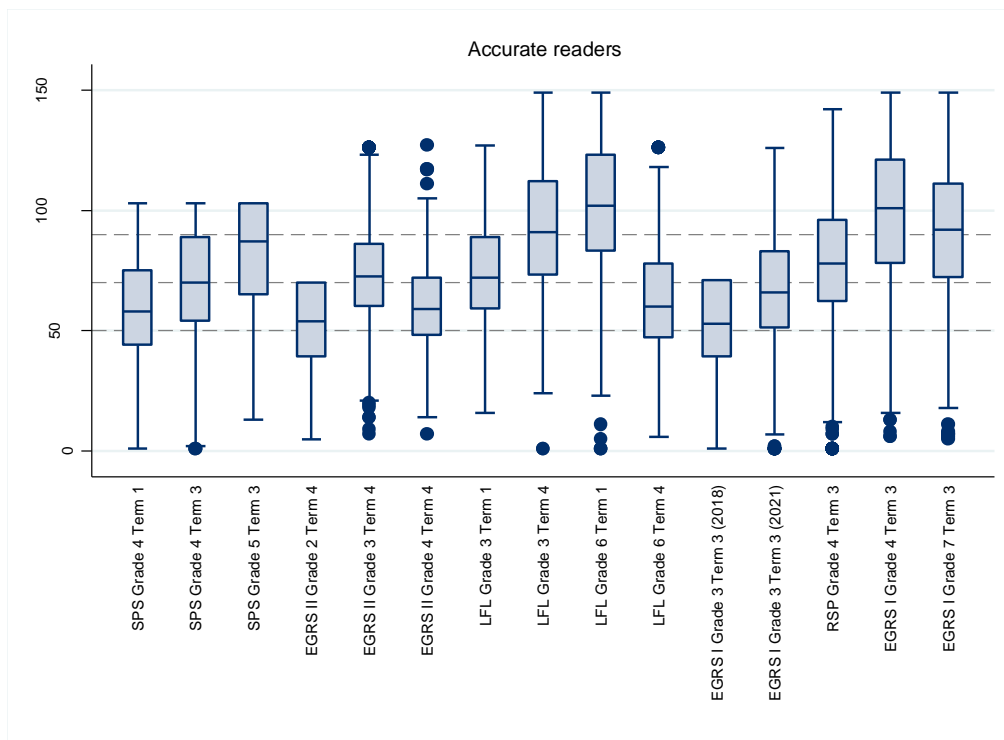
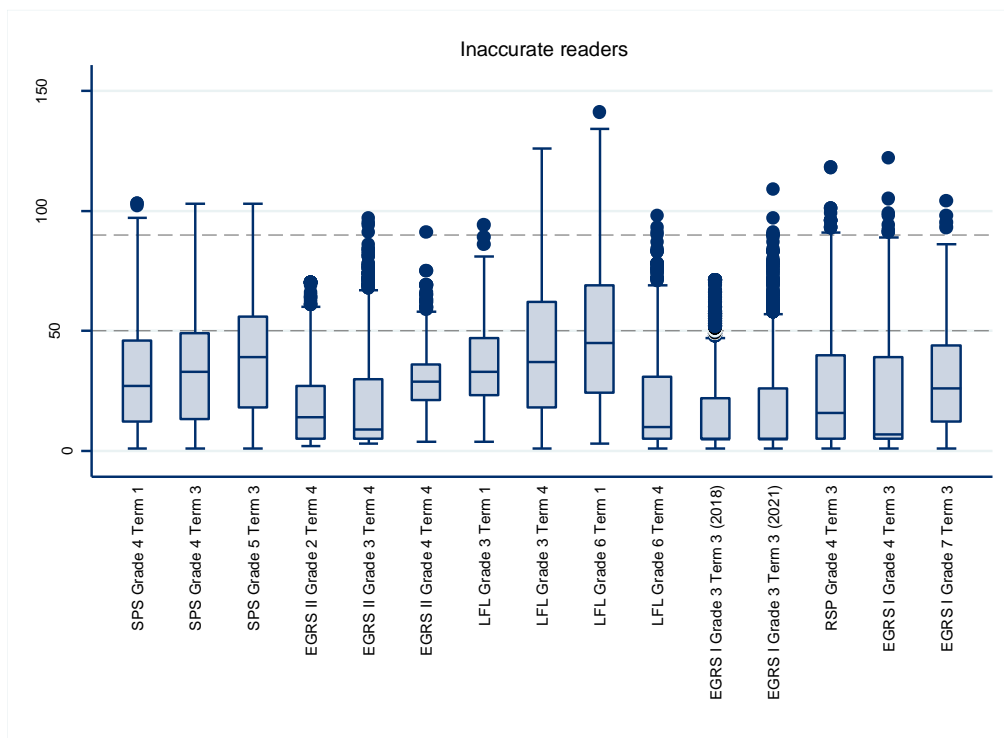


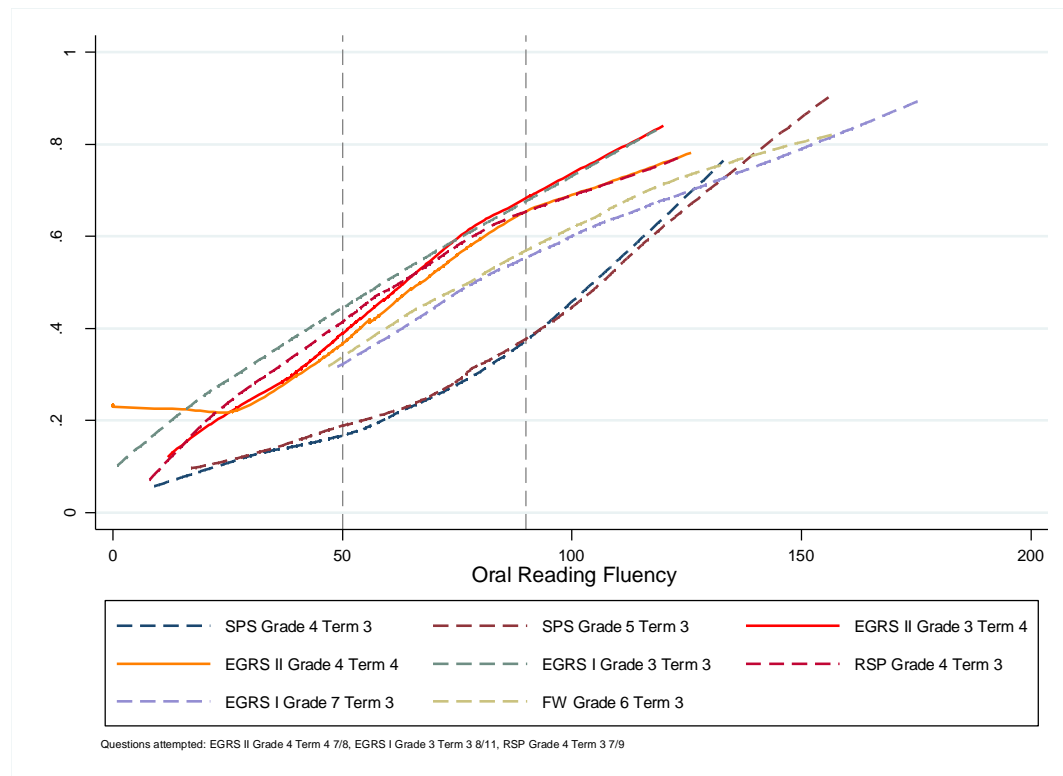
Figure 6: Speed distribution for learners reading with at less than 85% accuracy in EFAL



7.2 FLUENCY AND COMPREHENSION

We now examine the relationship between EFAL oral reading fluency²⁷ and oral reading comprehension. This relationship, in Figure 7, is only examined among learners who read far enough within the 3-minute time limit to attempt at least a subset of the comprehension questions (see Table 4). The average comprehension score at each level of fluency is plotted separately for all samples that meet these criteria. The figure includes reference lines at 50 and 90 correct words per minute.

Figure 7: Relationship between fluency and comprehension for samples completing a sub-set of all comprehension questions, EFAL

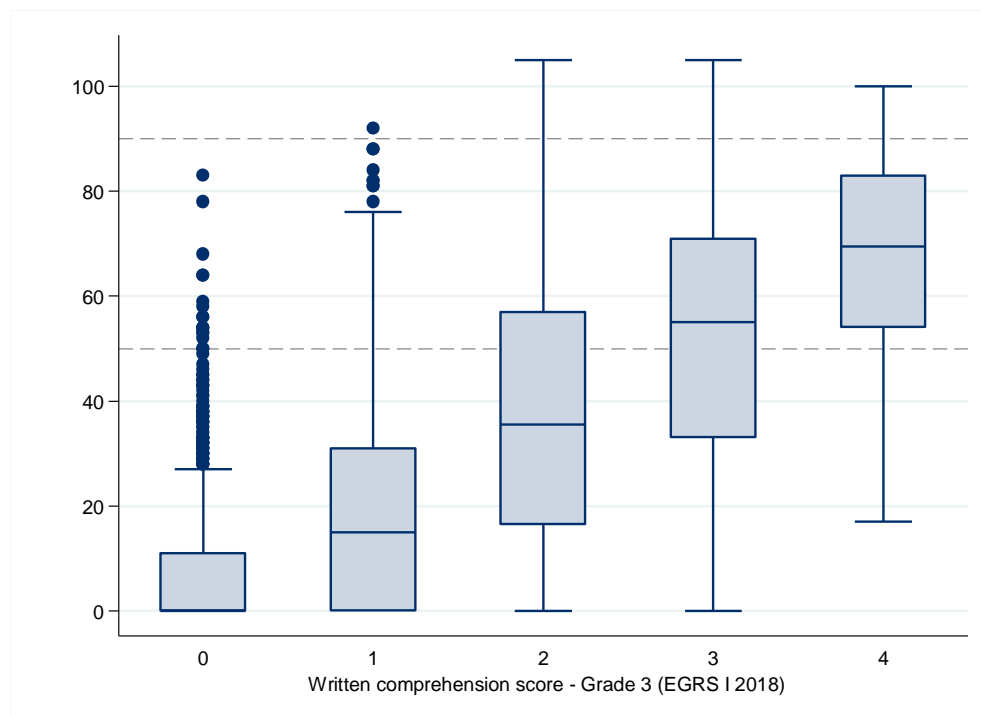


Although there are differences in the average comprehension level between samples, a positive fluency-comprehension relationship is observed. Across all eight samples, in the zone corresponding to where learners are reading inaccurately in English – i.e., getting less than 85 of 100 words attempted correct at reading speeds of less than 50 words per minute – their reading is not supportive of comprehension. Below 50 cwpm, learners score below or well below 40% on average in the ORF comprehension assessments. As reading fluency increases from 50 to 90 cwpm, in this accuracy zone, speeds are increasingly supportive of developing levels of comprehension. This is also confirmed when comparing box

²⁷ Fluency, in this context, means the number of correct words per minute from a passage of text

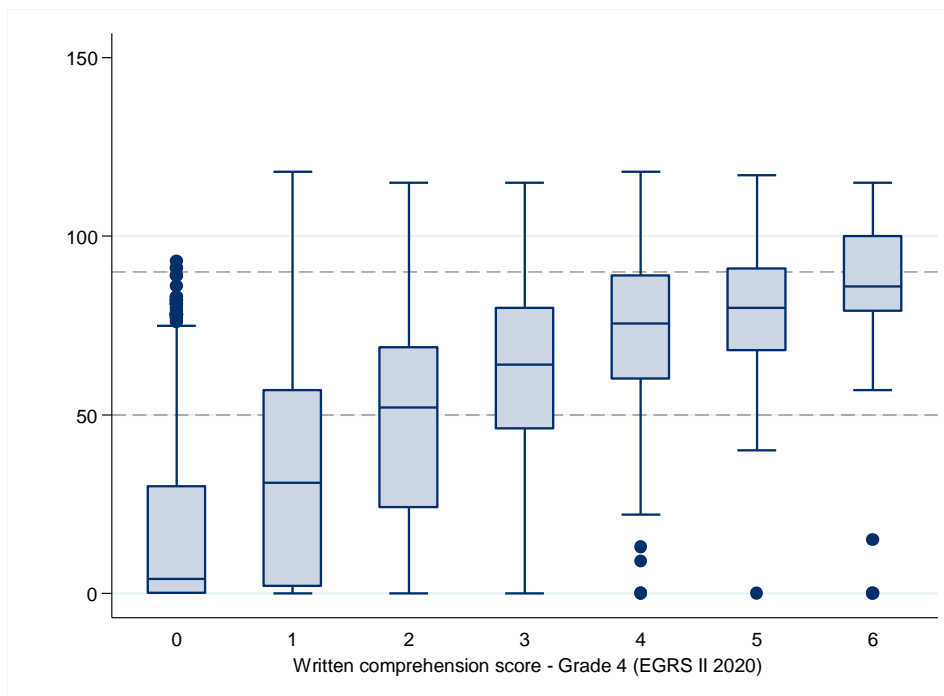
plots of EFAL fluency distributions by raw scores on an unrelated written comprehension assessment. Among Grade 3 and 4 samples, typically learners are only having success in written comprehension once reading at or above a threshold of 50 cwpm as seen in Figure 8 and Figure 9.

Figure 8: Grade 3 EFAL ORF distribution for a given written comprehension raw score (EGRS I 2018)



Depending on the study, and individual comprehension questions asked, the shape of this positive fluency-comprehension relationship varies. For example, for the SPS Grade 4 and 5 samples a *convex* fluency-comprehension relationship is identified over the range that overlaps with the other samples (seen in Figure 7), but initially at comparatively lower levels of comprehension for a given fluency level. Convexity is also observed when examining the fluency-comprehension relationships on most individual comprehension questions for these SPS samples (see Appendix Figure A 4 to Figure A 5). This implies that the fluency-comprehension relationship may be strongly influenced by differences in instructional practices across schools. Yet in the main, across 6 of 8 study samples considered, a concave relationship is observed implying diminishing returns to comprehension at higher levels of fluency. The identification of inflection points at which this flattening occurs ranges between 83 -123 cwpm across 6 studies.

Figure 9: Grade 4 EFAL ORF distribution for a given written comprehension raw score (EGRS II 2020)



For this reason, 90 cwpm seems to be a relevant yet simple metric for establishing a reading benchmark in EFAL. Accuracy is well developed at a speed corresponding to 90 words per minute as discussed earlier, and above 90 cwpm there are typically diminishing gains to comprehension of increasing reading fluency in English. We also observe from the box plots of fluency by raw ORF comprehension scores in Figure 10, that higher comprehension performance is typically only achieved by those reading at or above roughly 90 cwpm. Data from Grades 4 -7, used to plot ORF distributions against raw scores on an unrelated written comprehension test, also reveals similar patterns. In the Intermediate and Senior Phase, learners are only likely to be getting 75% or more for written EFAL comprehension tasks if their EFAL fluency is at or above 90 cwpm (see Figure 11 to Figure 13).

However, we caution that 90 cwpm does not signal a point at which comprehension skills have developed sufficiently and does not imply that fluency should not continue to improve. At 90 cwpm, EFAL ORF comprehension scores range on average between 40-70% (see Figure 7). Comprehension and language skills will still need to be taught, but once this milestone is reached, learners' fluency levels are in a zone where they are able to benefit further from language instruction in higher order skills.

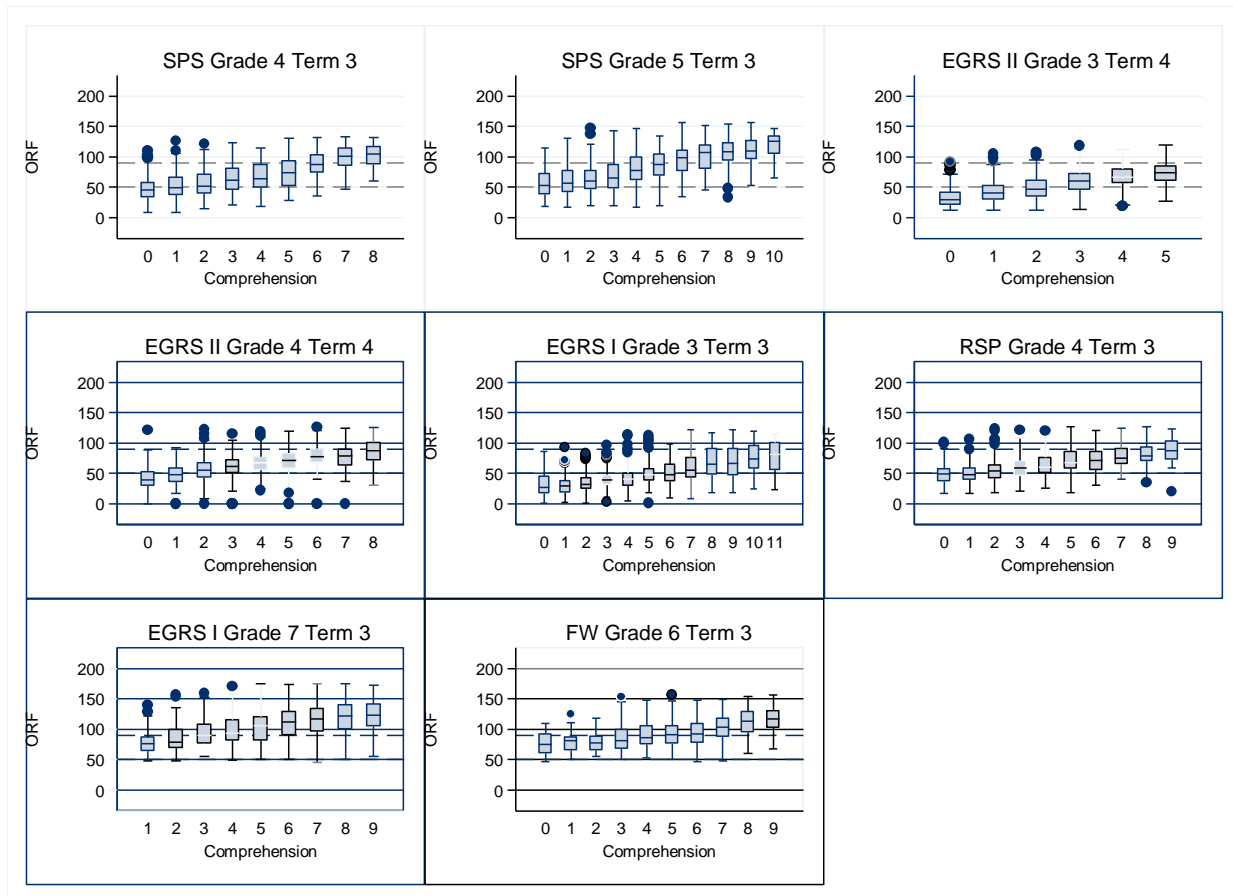
The regular patterns observed in examining the accuracy-speed and fluency-comprehension relationships in EFAL support the identification of:

Contract Number:, Order Number:

- a lower EFAL fluency threshold at around 50 cwpm, below which teaching should focus on improving decoding and fluency skills; and

a fluency benchmark at around 90 cwpm, above which teachers' attention should hone in on the strengthening of reading comprehension skills and vocabulary in EFAL, while continuing to support enhanced fluency.

Figure 10: EFAL ORF distribution for a given ORF comprehension raw score



Contract Number:, Order Number:

Figure 11: Grade 5 EFAL ORF distribution for a given written comprehension raw score (SPS 2019)

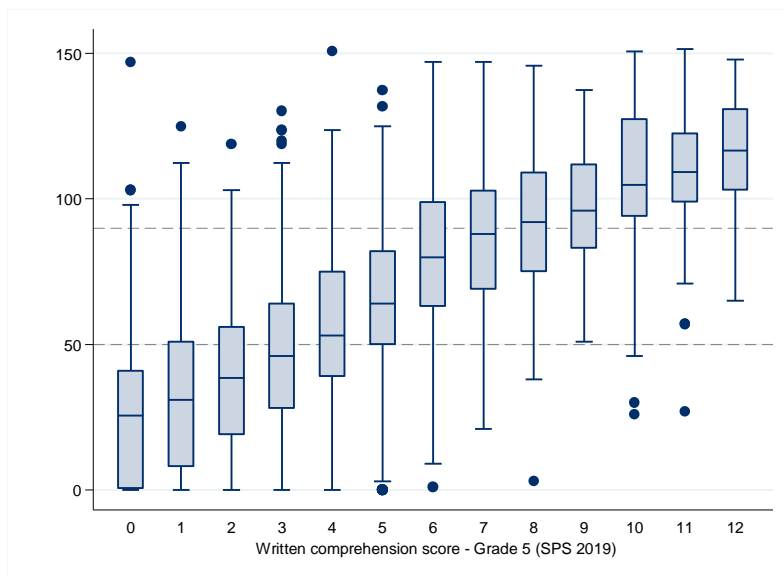


Figure 12: Grade 6 EFAL ORF distribution for a given written comprehension raw score (FW LP, 2021)

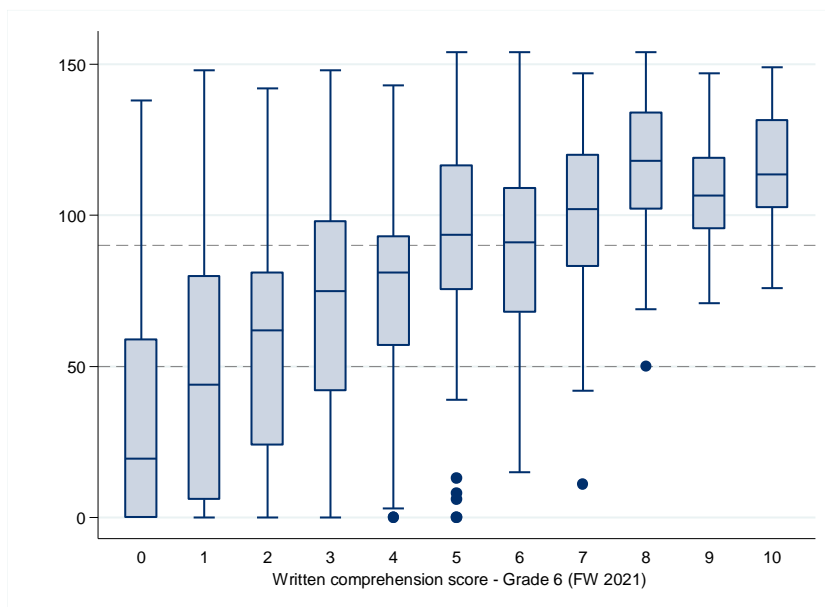
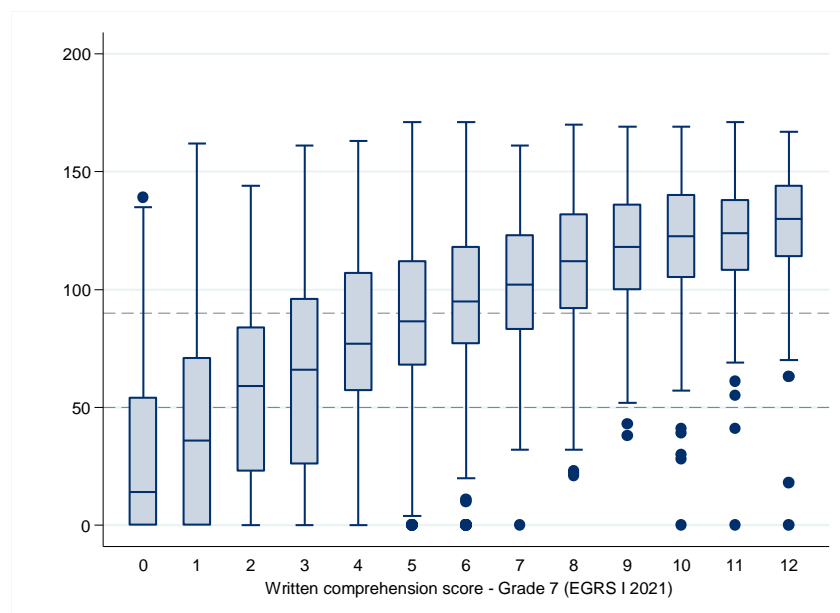


Figure 13: Grade 7 EFAL ORF distribution for a given written comprehension raw score (EGRS I, 2021)



7.3 PREDICTIVE VALIDITY OF EFAL FLUENCY THRESHOLD AND BENCHMARK

In this section, we explore whether an EFAL fluency threshold of 50 cwpm is predictive of improved fluency in later grades, particularly meeting the EFAL benchmark of 90 cwpm. We also consider how the threshold and benchmarks predict reading for meaning in later grades.

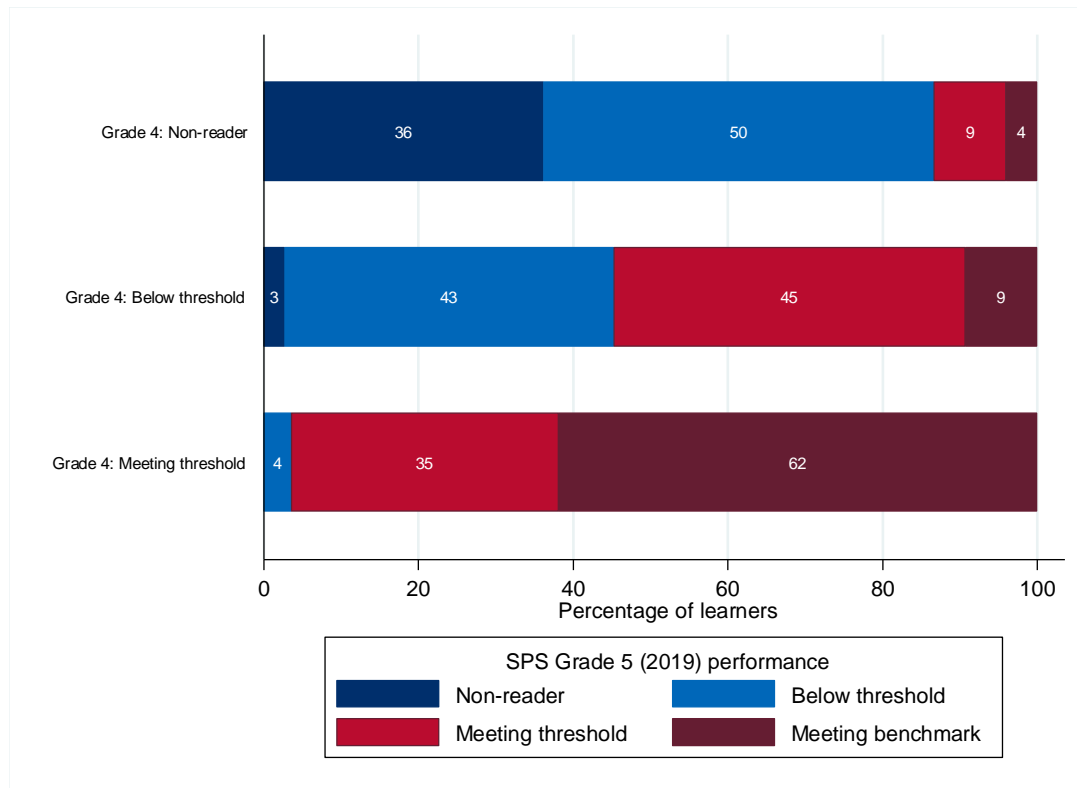
Figure 14 to Figure 16 summarise learner's fluency profiles in a later grade based on their fluency profile in an earlier grade. Among learner samples from the SPS study in the EC and KZN, the majority (62%) of those reading at or above the threshold of 50 cwpm by the end of Grade 4 reach the benchmark of 90 cwpm by the end of Grade 5 (see Figure 14). Among learner samples from the EGRS I study in NW, the vast majority (87%) of those reading at or above the threshold of 50 cwpm by the end of Grade 4 reach the benchmark of 90 cwpm by the end of Grade 7 (see Figure 15).

Despite the predictive validity of the threshold, school disruptions in 2020, where Grade 4s lost about 60% of a normal school year of improvement, resulted in slower reading development, even for those meeting the threshold. For example, examining the EFAL fluency profiles of EGRS II Grade 4s at the end of the first year of the pandemic, based on their fluency profiles a grade earlier, only 21% of those reaching the threshold by Grade 3 reach the benchmark of 90 cwpm by Grade 4 (see Figure 16). Three quarters of this group are reading between 50 and 90 cwpm. For those reading below the threshold of 50 cwpm at the end of Grade 3, none meet the benchmark at the end of Grade 4. Meeting the fluency threshold is

Contract Number:, Order Number:

a necessary but not entirely sufficient condition for meeting the benchmark. The pandemic situation highlights that decoding skills must be taught to high levels of mastery.²⁸

Figure 14: Fluency in Grade 5, by learners' fluency profile in Grade 4, EFAL



²⁸ It is noted that fluency declines are observed for some learners in the data across all samples. For example, learners reading at the threshold can be observed as reading nothing. This could reflect a combination of measurement error and learners just having a bad day on the day of the test.

Figure 15: Fluency in Grade 7, by learners' fluency profile in Grade 4, EFAL

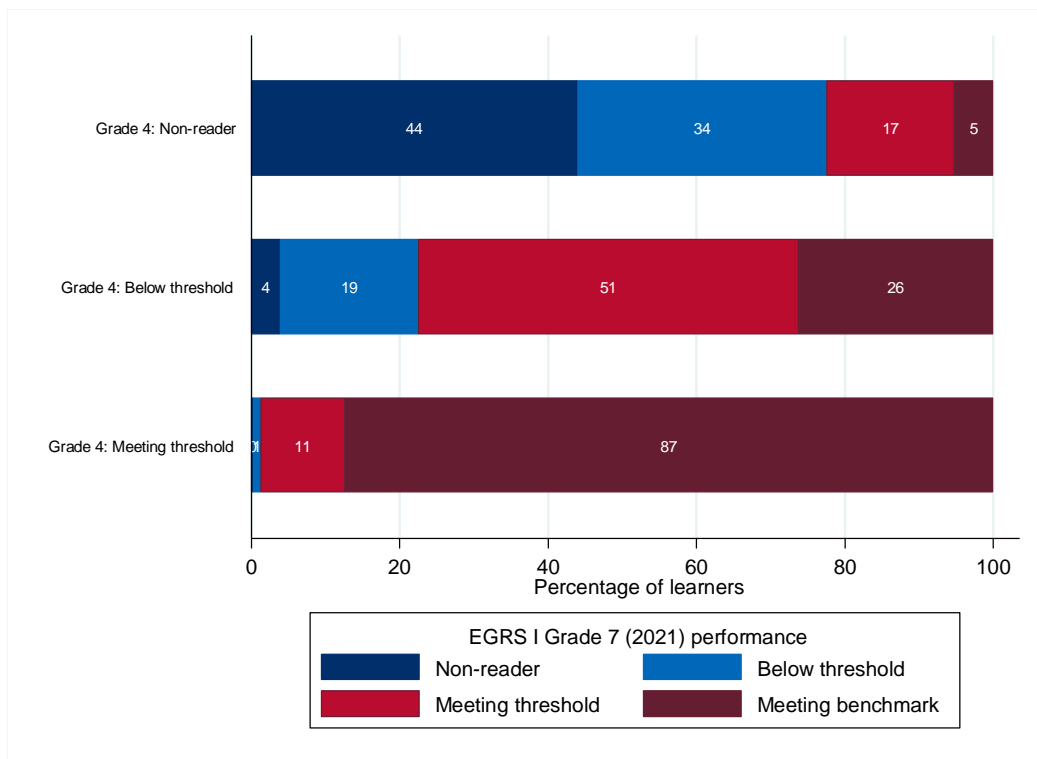
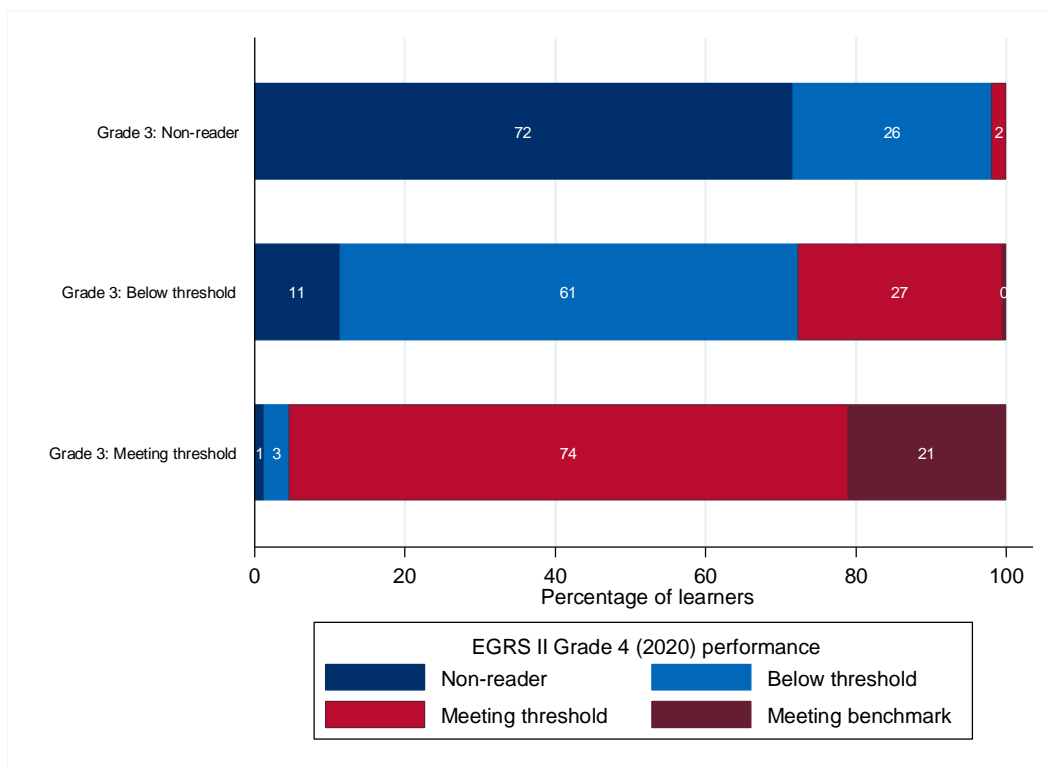


Figure 16: Fluency in Grade 4, by learners' fluency profile in Grade 3, EFAL



8 ALIGNING EFAL FLUENCY THRESHOLDS AND BENCHMARKS TO GRADE LEVELS

So far, the exploratory analysis is indicative of 50 cwpm as a lower fluency threshold and 90 cwpm as a benchmark in EFAL. But to what grade levels should the threshold and benchmark be aligned? To answer this, we consider at what grade levels they are currently attainable by our learner samples, balanced against EFAL fluency grade norms in other contexts and curriculum requirements.

8.1 EXAMINING THE ATTAINABILITY OF THE LOWER THRESHOLD AND BENCHMARK

To examine the attainability of the threshold/benchmark, Figure 17 categorises learners by study, grade and term into four fluency groups:

- unable to read one word (shown in dark blue),
- reading less than 50 cwpm (shown in light blue),
- meeting the threshold, reading 50 to 90 cwpm (shown in light red) and
- meeting the benchmark, reading at least 90 cwpm (shown in dark red).

Table 9 also revisits median EFAL fluency levels by grade. Although there are differences between studies, the general progression is clear when considering the results in this table with Figure 17:

- At the end of Grade 2 or the start of Grade 3, there are significant proportions of learners unable to read one word and few meet the threshold. At the median, samples are reading 11-15 cwpm (if non-readers are excluded, median fluency is 20-27 cwpm).
- By the end of Grade 3, most learners are reading at rates below the lower threshold of 50 cwpm (63-83% across samples); and 19-32% are still unable to read one word in English. Very few learners reach the benchmark of 90 cwpm (2-9% across samples) but about 19-37% reach the threshold. At the median, fluency ranges between 13-34 cwpm across samples including non-readers and 27-45 cwpm for samples excluding non-readers.
- By the end of Grade 4, 38-52% of samples were meeting the EFAL threshold of 50 cwpm. Just 8-15% meet the benchmark of 90 cwpm. Median fluency ranges from 21-46 cwpm, or 44-50 cwpm if non-readers are excluded from the samples.
- By the end of Grade 5, 29% of the rural no-fee learner sample in KZN or the EC meet the benchmark of 90 cwpm and median fluency is 62 cwpm with few non-readers.

- By the end of the Intermediate Phase (Grade 6), roughly 73-88% are meeting the EFAL threshold while 38-56% are meeting the benchmark. Median fluency ranges from 80-97 cwpm. By Grade 6, the benchmark is no longer aspirational but highly attainable.
- By the end of primary school (Grade 7), a large proportion of learners at 63% meet the EFAL benchmark of 90 cwpm, and 89% meet the threshold.

Given the slow development of reading in EFAL, the lower threshold of 50 cwpm is certainly aspirational, but not completely out of reach for learners in no-fee schools by the end of Grade 3. The benchmark of 90 cwpm is reached by large enough proportions of learners in Grade 6 -signalling the end of the Intermediate Phase - so that it is too low as a standard for Grade 6 and better positioned at the Grade 5 level. However, the next section clearly articulates how these are by no means high standards by ESL reading standards in other contexts.

Figure 17: Fluency profiles in EFAL by grade (excluding repeaters)

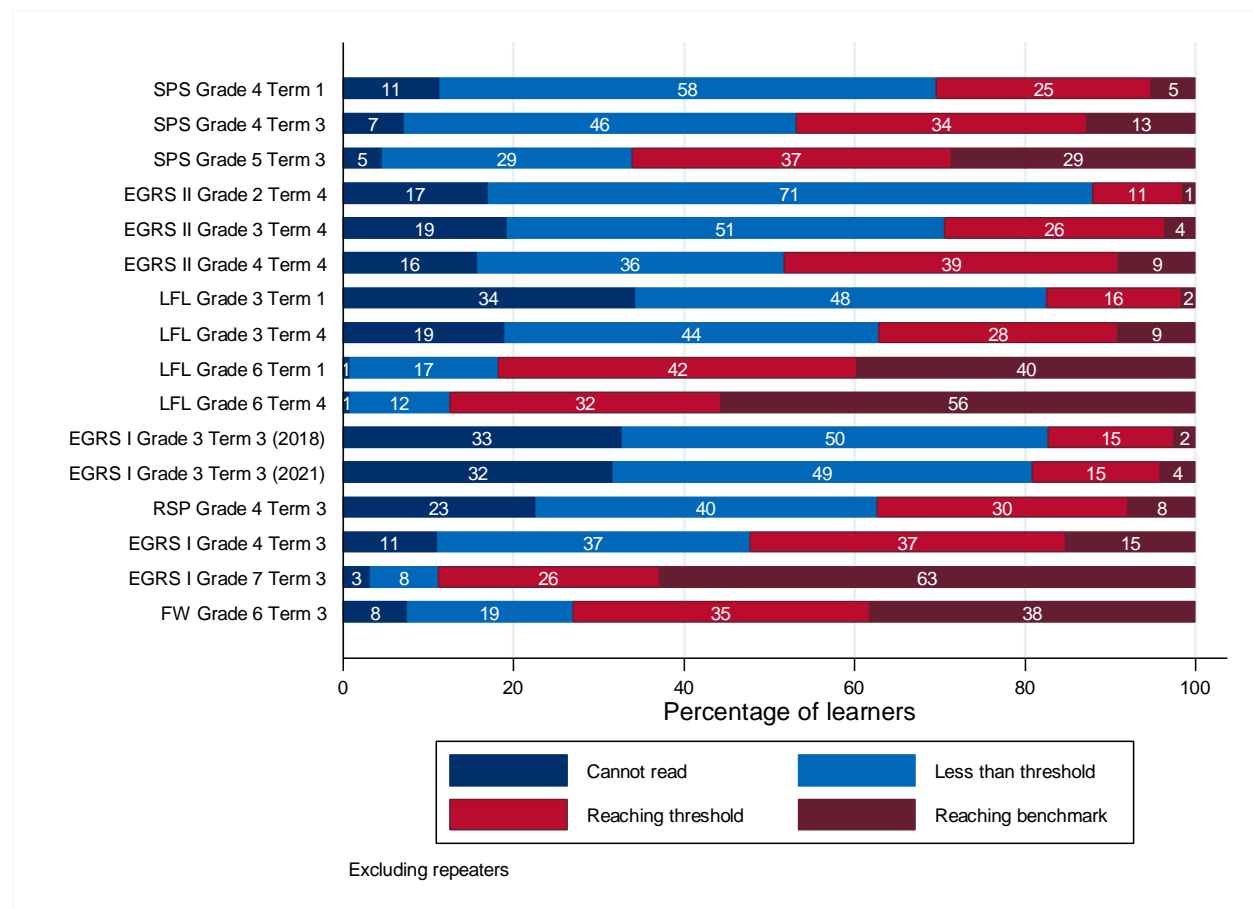


Table 9: Median EFAL fluency (50th percentile), South African learner samples (including repeaters)

Study	Grade	Term	Year	Text	Median ORF - All learners	Median ORF – learners with ORF > 0
EGRS II	2	IV	2018	b	11	20
LFL	3	I	2017	a	15	27
EGRS I	3	III	2018	a	13	27
EGRS I	3	III	2021	b	15	32
LFL	3	IV	2017	a	34	45
EGRS II	3	IV	2019	a	19	34
SPS Cohort II	4	I	2018	b	36	40
EGRS I	4	III	2018	a	30	48
SPS Cohort II	4	III	2019	b	46	49
RSP	4	III	2021	a	21	44
EGRS II	4	IV	2020	a	30	50
SPS Cohort II	5	III	2019	b	62	65
LFL	6	I	2017	c	80	81
LFL	6	IV	2017	c	97	97
FW LP	6	IV	2021	c	80	84
EGRS I	7	III	2021	c	89	95

8.2 EXAMINING THE LOWER THRESHOLD AND BENCHMARK AGAINST GRADE LEVEL ENGLISH FLUENCY NORMS FROM OTHER CONTEXTS

How does the threshold of 50 cwpm and benchmark of 90 cwpm align to suggested or actual fluency norms in ESL from other contexts?

Table 10 provides the Hasbrouck and Tindal (2017) fluency norms for HL English for the end of each primary grade for first language English speakers. We then add DIBELS ORF ranges for readers who are at minimal or negligible risk of reading failure (University of Oregon, 2017:123-124). Then subsequent columns summarise suggested norms for English second language speakers as discussed in Section 4.

At a minimum, we would like to aim for the South African EFAL threshold and benchmark to align with grades as per the Broward County BI level classification for ESL speakers (Psychological Services, Broward County Public Schools, 2021). These ‘BI’ learners are classified as intermediate English speakers, they communicate orally in English, mostly with simple phrases and/or sentence responses but make significant grammatical errors which interfere with understanding. A 50 cwpm threshold corresponds with the end of Grade 1 Broward BI norm, and 90 cwpm corresponds with the end of the Grade 3 Broward BI norm. Unfortunately, as discussed above, this would currently be far too aspirational for South African learners in no-fee schools. More realistically, a threshold of 50 cwpm at the Grade 3 level roughly aligns with the ESL Grade 3 median norm for Broward County group ‘AI’ students in the United States classified as ‘non-

Contract Number:, Order Number:

English speakers with minimal knowledge of English, demonstrating very little understanding, struggling to communicate meaning orally and unable to participate in regular classroom instruction.’

A benchmark of 90 cwpm aligns with Grade 5 norms for Broward A2 learners (Limited English speakers, demonstrating limited understanding and communicating orally in English with one- or two-word responses). Alternatively, it aligns with Grade 4 norms suggested by Jimerson et al. (2013).

Table 10: English fluency norms from international studies - End of grade median fluency (50th percentile) or in the case of DIBELS the ORF associated with being at negligible or minimum risk of reading failure at the end of each grade

Grade	L1: Hasbrouck & Tindal 50 th percentile (H&T) (2017)	L1: University of Oregon, DIBELS (2021) Minimal risk	L1: University of Oregon, DIBELS (2021) Negligible risk	L2: 70% of H&T (Anderson 1999)	L2: H&T less 25 words (Jimerson et al. 2013)	L2: Broward A1	L2: Broward A2	L2: Broward B1
1	60	39-75	76	42	35	43	43	46
2	100	94-127	128	70	75	44	52	69
3	112	114-135	136	78	87	49	74	89
4	133	125-158	159	93	108	65	80	103
5	146	137-156	157	102	121	85	89	104
6	146	141-159	160	102	121			
7	151 ^a	141-163	164	106				

Notes: *shading in light red – close to 90 cwpm benchmark, shading in light blue – close to 50 cwpm threshold. ^a This comes from the original Hasbrouck and Tindal (2006) norms. There are no updated norms for Grade 7 in Hasbrouck and Tindal (2017).

This comparison of norms leads us to the following conclusions about the grade levels aligned to the EFAL fluency threshold and benchmark.

- Weighing up South African no-fee school norms against international norms, it is most appropriate to align 90 cwpm to the end of Grade 5. This aligns with Broward A1 and A2 median reading norms at the end of Grade 5. Currently median EFAL fluency for Grade 5 learner samples in rural no-fee schools in KZN and EC is 62 cwpm. **The benchmark set at 90 cwpm at the end of Grade 5 is aspirational for learners in no-fee schools, yet internationally relevant.** Furthermore, if all learners are reading at least 90 cwpm by the end of Grade 5, they would easily be able to reach around 100 cwpm or more by the time they leave primary school which aligns with Anderson’s (1999) Grade 6 and 7 norms for ESL speakers.

- The new planned Systemic Evaluation²⁹, testing learner samples every year, will be aligned to Grades 3 and 6 in the primary phase or the end of the Foundation and Intermediate Phase (DBE, 2020a). Given current South African norms observed in Grade 6 samples, 90 cwpm is not aspirational enough as a benchmark for the end of the Intermediate Phase. **By aligning 100 cwpm to Grade 6, progress against this international ESL norm could be measured systematically if ORF assessments were added to planned systemic tests which are written assessments.**
- Ensuring all learners meet a threshold (of 50 cwpm) by the end of Grade 3 would by international standards be considered too easy. But just 17-37% of our learner samples reach the threshold of 50 cwpm by the end of Grade 3. It is therefore already an aspirational goal for the Grade 3 level in lieu of current South African norms, but attainable if reading skills were effectively taught and learners' EFAL reading development were on track by the end of Grade 2 (as discussed in the next sub-section).
- As a halfway point between the Grade 3 threshold and Grade 5 benchmark, **an intermediate EFAL fluency threshold of 70 cwpm is proposed for Grade 4.** It is not only a midpoint between the threshold and benchmark but internationally, a gain of 20 cwpm is the limit on improvements observed in existing median reading norms between the end of Grade 3 and Grade 4.
- Since the teaching of reading skills in EFAL is concentrated in the Foundation Phase, setting a threshold at the end of Grade 3 would not provide a sufficient guide to track reading development in earlier grades. A threshold guide is thus needed for Grade 2, as the teaching of reading in EFAL gains momentum in Grade 2. In the next section, **we propose a Grade 2 EFAL development threshold of 30 cwpm as a potential guide for teachers.** This a predictive milestone to meeting the threshold of 50 cwpm by the end of Grade 3.

In lieu of this international comparison, the EFAL grade-based thresholds and benchmarks established here – particularly for Grade 2 and 3 - should not be viewed as static over the long term. We recommend that

²⁹ Systemic assessments are assessments without any direct implications for individual learners, for instance with regard to grade promotion, but which help managers and planners understand the relative academic performance of whole schools, districts or provinces, or the country as a whole, and whether there is improvement at these levels.

as system improvements in the teaching of reading are observed, proposed EFAL fluency thresholds and benchmark should be shifted to earlier grades.

8.3 ESTABLISHING A GRADE 2 DEVELOPMENT FLUENCY THRESHOLD IN EFAL

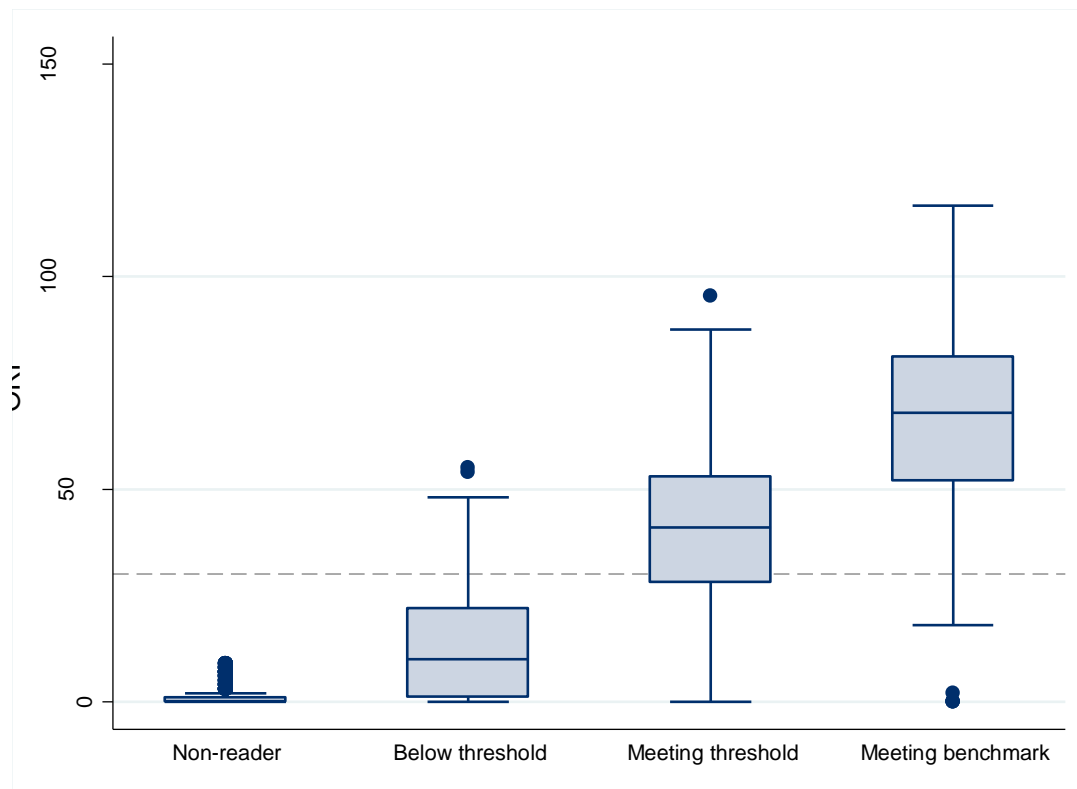
To establish a development fluency threshold in EFAL, we work backwards to Grade 2 from learners' Grade 3 fluency status using the EGRS II longitudinal data from Mpumalanga. We identify Grade 2 fluency levels by learner's ORF fluency category a year later in Grade 3 in Figure 18. We have excluded from the figure estimates, cases where measurement error is suspected.³⁰

Among learners meeting the Grade 3 threshold of 50 cwpm, there are very few reading less than 30 cwpm a year earlier at the end of Grade 2: 89% of Grade 2 learners reading between 1 and 29 cwpm do not meet the threshold of 50 cwpm by Grade 3. By contrast, 57% of those reading between 30 and 49 cwpm in Grade 2 reach the threshold of 50 cwpm by Grade 3.

In summary, **an EFAL fluency development threshold of 30 cwpm in Grade 2 is highly predictive of meeting the threshold of 50 cwpm by the end of Grade 3.**

³⁰ There were significant measurement issues at the grade 2 level. While passages were different and deteriorating performance is possible, all the very high ORFs for grade 2 have very low grade 3 ORF suggesting this is measurement error. We therefore remove learners whose grade 3 ORF is 10 or more words lower than their grade 2 ORF.

Figure 18: Grade 2 fluency by learner's Grade 3 fluency category a year later, EGRS II (Mpumalanga)



9 SUMMARY: EFAL FLUENCY THRESHOLD AND BENCHMARKS

When considered against proposed benchmarks or actual reading norms in English for second language English speakers from other international contexts, the fluency benchmark of 90 cwpm can be reasonably aligned with the Grade 5 level. We work forwards and backwards from this point to establish thresholds and benchmarks for the end of Foundation Phase and Intermediate Phase grades, starting with Grade 2 when EFAL learners are introduced to group guided reading and ending with Grade 6, corresponding with the last year of the Intermediate Phase. In summary, the following grade level development thresholds and benchmarks are proposed:

At the end of Grade 2, all EFAL learners should be reading grade-appropriate English passages with fluency levels at or above 30 cwpm.

- This should be viewed as a development threshold, a steppingstone to meeting the threshold of 50 cwpm by the end of Grade 3. This fluency level is far too low to support comprehension yet is predictive of whether learners can reach a fluency threshold of 50 cwpm a year later.

By the end of Grade 3, all EFAL learners should be reading grade-appropriate English passages at fluency levels at or above the EFAL threshold of 50 cwpm.

- When learners reach this minimum threshold of 50 cwpm, they start to derive a limited amount of meaning from what they are reading. 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 anything of what they have read, evidenced by very low oral comprehension or written comprehension scores.
- Meeting this threshold of 50 cwpm is strongly associated with whether learners will be reading at the EFAL benchmark of 90 cwpm in later primary grades. However, meeting the fluency threshold is a necessary but not entirely sufficient condition for meeting the benchmark. Decoding skills must be taught to mastery level, and learners should regularly practice reading to ensure fluency develops further.
- Moving towards all Grade 3 EFAL speakers reading at least at the threshold of 50 cwpm will require a refocus on teaching decoding skills in the Foundation Phase. Far too many children (19-33% across samples) are currently not able to read one word correctly in English by the end of Grade 3.

By the end of Grade 4, all EFAL learners should be reaching 70 cwpm when reading grade-appropriate English passages.

- This a steppingstone to reaching the EFAL benchmark of 90 cwpm by the end of Grade 5. It is not only a midpoint between the threshold and benchmark but internationally, 20 cwpm is the limit on annual improvements in median reading norms observed between the end of Grade 3 and Grade 4.
- Achieving a situation where all learners are reading at least 70 cwpm by the end of Grade 4 will require significant increases in current fluency rates.

By the end of Grade 5, all EFAL learners should be reading grade-appropriate passages at or above the EFAL benchmark of 90 cwpm.

- Reading at this fluency level is necessary to be able to reach higher levels of comprehension. Only learners reading at or above 90 cwpm are getting most questions correct about the passage they have read.

Contract Number:, Order Number:

- However, fluency levels at or above 90 cwpm are necessary but not sufficient to support improvements in comprehension.
- This milestone signals the point at which teachers' focus should shift to the teaching of the skills and strategies learners need to tackle written comprehension while encouraging vocabulary and language development.

By the end of Grade 6, all EFAL learners should be reaching 100 cwpm when reading grade-appropriate English passages.

- In international studies, reading around 100 cwpm is suggested as the 50th ORF percentile for Grade 5 second language English speakers. It is therefore grade appropriate by international standards.

Reaching these EFAL thresholds and benchmarks should not be viewed in isolation of reaching home language reading thresholds and benchmarks. The development of decoding skills in African home language reading provides an important foundation for learning to read in English as African languages and English are alphabetic languages. The CAPS EFAL curriculum develops from the assumption that when children begin to read and write in their additional language, they already know how to decode in their home language. It assumes that they have grasped concepts of print and have prior knowledge of sound-spelling relationships.

Although we have not specified a letter-sound benchmark for EFAL, 40 correct letter sounds per minute at the end of Grade 1 will also suffice. However, consonants and vowels across home language and EFAL will not be sounded in exactly the same way. Consequently, starting in Grade 1 and into Grade 2, children need to learn where sound-spelling relationships are different in their home and additional languages. For example, 'th' in English represents two different voiced and voiceless fricative sounds, which are different to the aspirated sound which 'th' represents in African languages. English vowels can also be challenging for African language speakers, and this is made more difficult by the variety of ways in which these vowels are spelt. As CAPS acknowledges, children will require practice in applying decoding knowledge in home language to text in English (DBE 2011: 15).

10 RECOMMENDATIONS

Although this research advances the establishment of reading benchmarks and thresholds in South African languages, the true value of this body of research in supporting policy and practice will only be realised when linked to a national programme to assess and monitor early grade reading skills (Ardington et al., 2021a:14). If national EGRA-type assessments were introduced, the threshold and benchmarks could be used to monitor sector progress in reading in the early grades. Currently, early grade reading is not being systematically measured in South Africa at district, provincial or national levels. Without credible measurement of foundational early reading skills, it will not be possible to track reading progress. This leads us to key recommendations.

1. Implement a national system to test and monitor early grade reading skills:

Measurement of early reading skills would bring into focus the importance of early reading skills, reshaping policy priorities towards developing reading competency and directing resources to promote reading development. Written assessments of reading comprehension, typically used in primary school testing, are unable to detect foundational reading skills. The proposed Systematic Evaluations – a sample-based testing system (DBE, 2020a) – could be bolstered to include oral reading fluency assessments in addition to written assessments. At the primary grades, the Systemic Evaluation tests will be aimed annually at the Grade 3 and 6 level. Established thresholds and benchmarks can be applied at these grade levels to identify and track over time what proportion of learners can meet them and to remediate learners who do not meet them.

2. Continue to collect early grade reading data to establish threshold and benchmarks in all South African languages:

Thresholds and benchmarks for reading are only likely to be linked to a national testing system if they are available in all official South African languages. Early grade reading data collection exercises will need to be repeated in other languages as identified in Table 11 below. It would also be prudent to re-evaluate current benchmarks and thresholds as more data is made available, with different samples from different provinces and in lieu of the possibility that reading norms can shift over time.

Table 11: Current status of DBE donor/partner planned early grade reading benchmarking in all official languages

Language	Language group	Establishing reading benchmarks: current status
Sepedi Sesotho Setswana	Sotho-Tswana languages	in progress - establish whether Setswana benchmarks are applicable available - establish whether Setswana benchmarks are applicable complete (current report)
Siswati isiNdebele isiZulu isiXhosa	Nguni languages	complete data not available – but Nguni language benchmarks assumed complete complete
Xitsonga Tshivenda Afrikaans English (Additional language) English (home language)		data collection planned for August 2022 data collection planned for 2023 data collection in progress complete (current report) norms available from international contexts

3. **All forthcoming early grade reading data collection programs should ensure evidence-based best practices are followed so that new assessments can support benchmarking exercises.** Organisations considering collecting early grade reading assessment data for independent research projects or evaluations could collaborate to support national reading benchmarking exercises. Their data could be used for benchmarking purposes if appropriate protocols and processes are followed. This includes applying a 3-minute time allocation rule to ORF assessments. The EGRS I / RSP process of instrument development, with multiple rounds of testing and suitable protocols to support the collection of assessments of decoding skills provides a best practice scenario for how this should be done. The collaborative work of linguists, data analysts, home language specialists and DBE officials was critical throughout this process, and in establishing the final benchmarks/thresholds.

4. **Ensure that EGRA-type assessments are included as a critical aspect of formative assessment practice in primary schools:** EGRA-type assessments should be ongoing in all primary schools. These EGRA-assessments are not only relevant for the Foundation Phase grades but should be applied at the Intermediate Phase level, especially given the devastating learning losses associated with COVID-19 related school disruptions (Ardington et al., 2021b). Our analysis at the Grade 7 level was indicative of significant percentages of learners leaving primary school not having met the grade fluency benchmark of 60 cwpm in Setswana (13-48%) and Grade 5 fluency benchmark of 90 cwpm in EFAL (37%). The progressive roll-out of EGRA training for teachers by the Department of Basic Education in 2015 should be leveraged to promote EGRA testing, equipping teachers with the skills to administer these assessments (Maboya, 2020). The newly established benchmarks should be connected this, as they provide a framework for teachers to easily interpret the results from the EGRA assessment. As teachers conduct

EGRA-assessments, guided by thresholds and benchmarks, this will help them to identify early-on whether learners are on track, and align their instructional practice with each learner's level of reading development.

5. Preservice Initial Teacher Education programs should also reflect familiarity with EGRA assessment procedures: To support the inclusion of EGRA-type assessments in formative assessment and its effective implementation in schools, new teachers entering the system need to be equipped to assess early reading development and identify and remediate struggling readers through appropriate preservice Initial Teacher Education programs.

6. Resources to promote reading development through a multi-sectoral approach need to be consistently prioritised in policy and in budgets: Key resources required to promote reading development and instil an enjoyment of reading include training teachers in how to teach reading effectively through improved pre-service and in-service training; and ensuring that learners have sufficient and suitable books to read. Relatedly, the growing problem of large class sizes in the early grades also needs to be addressed as individualised reading instruction or even assessment is hindered when class sizes exceed prescribed recommendations (DBE, 2020a:106). More effort also needs to be given to considering how reading can receive higher priority in homes so that children are more exposed to oral language and print at earlier ages. Children are entering school with underdeveloped emergent literacy and language skills (Dawes et al., 2017), making the work of teachers and acquiring decoding skills much harder. There is also little evidence to suggest that decoding skills are being introduced effectively in Grade R.

7. Urgent allocation of resources for large-scale reading remediation programs in all schools: Finally, resources urgently need to be allocated so that reading remediation programs are available in all schools, and that schools are equipped with enough personnel and resources to support not just a few learners, but many. Teacher assistant support models have proven to be effective in initial pilot projects (Ardington & Henry, 2021), and could be further experimented with as a working model to support reading remediation at scale. Effective remediation is particularly necessary in a context where COVID-19 schooling disruptions have impacted severely on learner's reading development, particularly in the early grades (Ardington et al., 2021b).

REFERENCES

- Abadzi, H., 2012. Developing Cross-Language Metrics for Reading Fluency Measurement, Developing Cross-Language Metrics for Reading Fluency Measurement. <https://doi.org/10.1596/26819>
- Alcock, KJ, Ngorosho, D, Deus, C & Jukes, MCH. 2010. We don't have language at our house: Disentangling the relationship between phonological awareness, schooling and literacy. *British Journal of Educational Psychology*, 80(1), 55-76. Doi:10.1348/000709909X424411
- Anderson, N.J. 1999. Improving reading speed: Activities for the classroom. *English Teaching Forum* 37: 2–5.
- Ardington, C. & Henry, J. 2021. Funda Wande Limpopo Impact Evaluation. Midline Report. SALDRU, University of Cape Town. Cape Town.
- Ardington, C., Wills, G., Pretorius, E., Deghaye, A., Mohohlwane, & N., Menendez. 2021a. Benchmarking oral reading fluency in the early grades in Nguni languages. *International Journal of Educational Development*, 81, 102433. DOI: 10.1016/j.ijedudev.2021.102433
- 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.; Wills, G. & Kotze, J. 2021b. COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, vol 86, 102480.
- Aro, M. 2017. Learning to read Finnish. In Verhoeven, L & Perfetti, C. (eds), *Learning to read across languages and writing systems* (pp 416-435). Cambridge, Cambridge University Press.
- Betts, E. A. 1946. *Foundations of reading instruction, with emphasis on differentiated guidance*. American Book Co.
- Castles, A., Rastle, K. & Nation, K., 2018, 'Ending the reading wars: Reading acquisition from novice to expert', *Psychological Science in the Public Interest* 19(1), 5–51. <https://doi.org/10.1177/1529100618772271>
- Dale, E. & Chall, JS. 1948. A formula for predicting readability and A formula for predicting readability: Instructions. *Educational Research Bulletin*, XXVII (January 21 and February 18, 1948), 11-54.
- Dawes, A., Biersteker, L., Girdwood, E., Snelling, M. & Tredoux, C. 2017. Innovation Edge Briefing Document. The Development and Age Validation of the Early Learning Outcomes Measure (ELOM). Available: <http://elom.org.za/wp-content/uploads/2017/07/ELOM-Briefing-Document.pdf>
- Deno, SL, Fuchs, LS, Marston, D & Shin, J. 2001. Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, 30(4), 507–524.
- Department of Education and Training. 1988. Setswana. Terminology and orthography, no. 4. Government Printer Pretoria, South Africa.
- DBE, 2011. Curriculum and Assessment Policy Statement. Foundation Phase Grades 1-3. English First Additional Language. Department of Basic Education, South Africa. Pretoria. <https://www.education.gov.za/Portals/0/CD/National%20Curriculum%20Statements%20and%20Vocational/CAPS%20ENGLISH%20FAL%20GR%201-3%20FS.pdf?ver=2015-01-27-155321-957>
- DBE, 2017. Summary Report: Results of Year 2 Impact Evaluation. The Early Grade Reading Study (EGRS). Department of Basic Education, South Africa. Pretoria. <https://www.education.gov.za/Portals/0/Documents/Reports/EGRS%20Summary%20Report.pdf?ver=2017-08-17-090215-583>

Department of Basic Education and the University of the Witwatersrand. 2020. Early Grade Reading Study 2017-2019, Waves 1-4 Merged [dataset]. Version 1. Pretoria: DBE and Wits [producers], 2020. Cape Town: DataFirst [distributor], 2020. DOI: <https://doi.org/10.25828/qwx3-4m77>

DBE, 2020a. Action Plan to 2024. Towards the realisation of Schooling 2030. August 2020. Department of Basic Education. Pretoria, South Africa.

DBE, 2020b. National Framework for the Teaching of Reading in African Languages in the Foundation Phase. Department of Basic Education. Pretoria, South Africa.

Dowd, A.J., Bartlett, L., 2019. The Need for Speed: Interrogating the Dominance of Oral Reading Fluency in International Reading Efforts. *Comp. Educ. Rev.* 63, 189–212. <https://doi.org/10.1086/702612>

Durgunoğlu, A & Öney, B. 1999. A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing: An Interdisciplinary Journal*, 11, 281–299.

EGRS II. 2018. Year 2 report: Learner performance after the second year of implementation. Pretoria: Department of Basic Education.

Florit, E & Cain, K. 2011. The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, 23, 553–576. doi:10.1007/s10648-011-9175-6

Fuchs, LS, Fuchs, D, Hosp, MK & Jenkins, JR. 2001. Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5(3), 239–256.

Geva, E. & Zadeh, Z.Y. 2006. Reading efficiency in native English-speaking and English-as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading* 10(1): 31–57.

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. *Read. Teach.* 59, 636–644. <https://doi.org/10.1598/RT.59.7.3>

Hasbrouck, J. & Tindal, G., 2017. An update to compiled ORF norms (Technical Report No. 1702). Eugene, OR, Behavioral Research and Teaching, University of Oregon.

Hoover, WA & Gough, PB. 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., Tshela, 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.

Jimerson, S. R., Hong, S., Stage, S., & Gerber, M. 2013. Examining oral reading fluency trajectories among English language learners and English speaking students. *New Approaches in Educational Research*, 2(1), 3–11.

Jukes, M., Pretorius, E. Schaefer, M. Tjasink, K. Roper, M. Bisgard, J. Mabhena, N. 2020. Setting Reading Benchmarks in South Africa. Khulisa Management Services: Johannesburg. Department of Basic Education: Pretoria. Available from: https://pdf.usaid.gov/pdf_docs/PA00X1NZ.pdf.

Katz, J., 2020. Back to basics - comparing the orthographic, phonic and grammatical features of English and African languages to improve literacy teaching. Presentation at PrimTED Literacy Working Group Seminar. Materials for literacy teacher programmes.

Katz, L., Frost, R., 1992. The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis, in: Frost, R., Katz, L. (Eds.), *Orthography, Phonology, Morphology, and Meaning*. Elsevier Science Publishers, Amsterdam, pp. 67–84. [https://doi.org/10.1016/S0166-4115\(08\)62789-2](https://doi.org/10.1016/S0166-4115(08)62789-2)

- Kendeou, P, Papadopoulos, TC & Kotzapoulou, M. (2013). Evidence for the early emergence of the simple view of reading in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal*, 26, 189–204. doi:10.1007/s11145-012-9361-z
- Kaminski RA & Good RH. III. 1996. Toward a technology for assessing basic early literacy skills. *School Psychology Review*, 25(2), 215-227.
- Kim, Y.-S., Park, C., & Wagner, R.K. 2014. Is oral/text reading fluency a “bridge” to reading comprehension? *Reading and Writing*, 27(1), 79–99. doi:10.1007/s11145-013- 9434-7
- Kim, YG. 2017. Why the simple view of reading is not simplistic: Unpacking the simple view of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading*, 21, 310–333. <http://dx.doi.org/10.1080/10888438.2017.1291643>
- Kim, YG. (2020). Hierarchical and Dynamic Relations of Language and Cognitive Skills to Reading Comprehension: Testing the Direct and Indirect Effects Model of Reading (DIER). *Journal of Educational Psychology*, 112(4), 667-684. <http://dx.doi.org/10.1037/edu0000407>
- Kilpatrick, DA 2012. *Essentials of assessing, preventing and overcoming reading difficulties*. Hoboken, NJ: John Wiley & Sons.
- Land, S. 2015. *Reading isiZulu: Reading processes in an agglutinative language with a transparent orthography*. Unpublished doctoral thesis, University of KwaZulu-Natal, Pietermaritzburg, South Africa.
- LaBerge, D., Samuels, S.J., 1974. Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- Lekgoko, O. & Winskel, H. 2008. Learning to read in Setswana and English: cross-language transference of letter-knowledge, phonological awareness and word reading skills. *Perspectives in Education*, 26(4): 57 - 73.
- Leppänen, I, Aunola, K, Niemi, P & Nurmi, J-E. (2008). Letter knowledge predicts Grade 4 reading fluency and reading comprehension. *Learning and Instruction*, 18, 548-564.
- Lesaux, N. K. and L. S. Siegel (2003). The development of reading in children who speak English as a second language. *Developmental Psychology* 39 (6): 1005-1019.
- Lipka, O. & Siegel, L.S. 2007. The development of reading skills in children with English as a second language. *Scientific Studies of Reading* 11(2): 105–131.
- Liswaniso, B.L. 2021. *The design and effects of a catch-up reading intervention for Grade 5 teachers and learners in Namibia*. Unpublished PhD thesis, University of South Africa.
- Maboya, M., 2020. *Report on Foundational Skills of Literacy and Numeracy*. Curriculum Policy, Support and Monitoring, Department of Basic Education, Pretoria, South Africa.
- Malda, M, Nel, C & van de Vijver, FJR. 2014. The road to reading for South African learners: The role of orthographic depth. *Learning and Individual Differences*, 30: 34-45. <http://dx.doi.org/10.1016/j.lindif.2013.11.008>
- Makaure, ZP. 2021. *The contribution of phonological processing skills to early literacy development in Northern Sotho-English bilingual children: A longitudinal study*. Unpublished doctoral thesis, University of South Africa, Pretoria.
- Melawana ya Mkwalo le Mopeleto SETSWANA 2008. PanSALB. Arcadia. Ramagoshi, R.M. 2020. *A re Bueng Setswana*.
- Melby-Lervåg, M, Lyster, S & Hulme, C. 2012. Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, 138(2):322-52. doi: 10.1037/a0026744.
- Menendez, A & Ardington, C. 2018. *Impact Evaluation of USAID/South Africa Story Powered School Programme – Baseline*.

Contract Number:, Order Number:

Mutema, F. 2021. The development of reading literacy skills in the early years of primary schooling: a case of four Zimbabwean schools. Unpublished doctoral thesis, University of South Africa, Pretoria.

Piper, B. 2009. Integrated Education Program: Impact Study of SMRS Using Early Grade Reading Assessment in Three Provinces in South Africa. RTI International. Research Triangle Park, NC.

Piper, B., & Zuilkowski, SS. 2015. Assessing reading fluency in Kenya: Oral or silent assessment? *International Review of Education*, 61(2): 153-171. DOI 10.1007/s11159-015-9470-4

Pretorius, EJ. 2014. Supporting transition or playing catch up in Grade 4? Implications for standards in education and training. *Perspectives in Education* 32(1): 51-76.

Pretorius, E.J., Spaul, N., 2016. Exploring relationships between oral reading fluency and reading comprehension amongst English second language readers in South Africa. *Read. Writ.* 29, 1449–1471.
<https://doi.org/10.1007/s11145-016-9645-9>

Pretorius, EJ. 2018. Getting it right from the start: some cautionary notes for early reading instruction in African languages. In N Spaul & JP Comings (Eds). *Improving Early Literacy Outcomes: Curriculum, Teaching and Assessment* (pp63-80). Leiden: IBE/BRIL

Psychological Services, Broward County Public Schools. 2012. Oral Reading Fluency Data for English Language Learners (ELL). August 2012.

Room to Read, 2018. Setting Data-Driven Oral Reading Fluency Benchmarks Guidance Note. Research Triangle Park, NC.

RTI International. 2017. All Children Reading-Asia: EGRA Benchmarks and Standards Research Report

Samuels, S.J., Flor, R.F., 1997. The importance of automaticity for developing expertise in reading. *Read. Writ. Q.*
<https://doi.org/10.1080/1057356970130202>

Seymour, PHK, Aro, M & Erskine, JM. 2003. Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143–174.

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

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

Spaul, N., Pretorius, E. & Mohohlwane, N., 2020, 'Investigating the comprehension iceberg: Developing empirical benchmarks for early-grade reading in agglutinating African languages', *South African Journal of Childhood Education* 10(1), a773. <https://doi.org/10.4102/sajce.v10i1.773>

Stanovich, KE. 2000. *Progress in Understanding reading: Scientific foundations and new frontiers*. New York: The Guildford Press.

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

Torppa, Georgiou, Lerkkanen, Niemi, Poikkeus, Nurmi, 2016. Examining the Simple View of Reading in a Transparent Orthography: A Longitudinal Study From Kindergarten to Grade 3. *Merrill. Palmer. Q.* 62, 179.
<https://doi.org/10.13110/merrpalmquar1982.62.2.0179>

University of Oregon. 2021. 8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Administration and Scoring Guide, 2021 Edition. Eugene, OR: University of Oregon.
https://dibels.uoregon.edu/sites/dibels1.uoregon.edu/files/2021-10/UE_Dibels8_Scoring_Guide_100121.pdf

Wang, Z., Sabatini, J., O'Reilly, T., Weeks, J., 2019. Decoding and Reading Comprehension: A Test of the Decoding Threshold Hypothesis. *J. Educ. Psychol.* 111, 387–401.

Contract Number:, Order Number:

Wills, G. & van der Berg, S. 2020. Measuring School Leadership and Management and Linkages with Literacy: Evidence from rural and township primary schools in South Africa. *Educational Management Administration and Leadership*. 45 (5) 708-731. DOI: 10.1177/1741143220915923

Zenex Literacy Project: Learner reading assessment results. 2017. ERA: Stellenbosch University.

APPENDIX

Instrument development: Pilot samples

Piloting of the learner instruments was conducted in three rounds. Table A1 summarises the sample size by grade for each round of piloting. The first pilot was very small with a target of around 30 learners in each grade and was conducted in two schools in Gauteng. Unfortunately, this meant that around half the assessed learners were not Setswana home language speakers. We decided to include all learners in the analysis for the first pilot as the non-Setswana home language learners actually performed slightly better than the Setswana home language learners and we were concerned about further reductions to the already small sample. The second and third pilots were larger and were conducted in schools in the North-West province. The percentage of Setswana home language speakers increased to between 88% and 95% of the samples for each grade. Given the purpose of the pilot, the assessments were suspended if learners were unable to read at least one word of the first passage. This allowed more time to assess learners who were able to read. Although 44% of the Grade 3 learners in pilot 2 were unable to read a word, the fieldwork team were still able to assess 62 learners who could read

Table A 1. Pilot samples

	Pilot 1			Pilot 2			Pilot 3		
	Learners	% Setswana home language	% non-readers	Learners	% Setswana home language	% non-readers	Learners	% Setswana home language	% non-readers
Grade 3	30	50%	3%	111	88%	44%	125	93%	16%
Grade 4	27	52%	10%	74	91%	20%	105	89%	11%
Grade 7	29	45%	0%	85	95%	8%	95	95%	1%

EFAL benchmarking results: Additional tables and figures

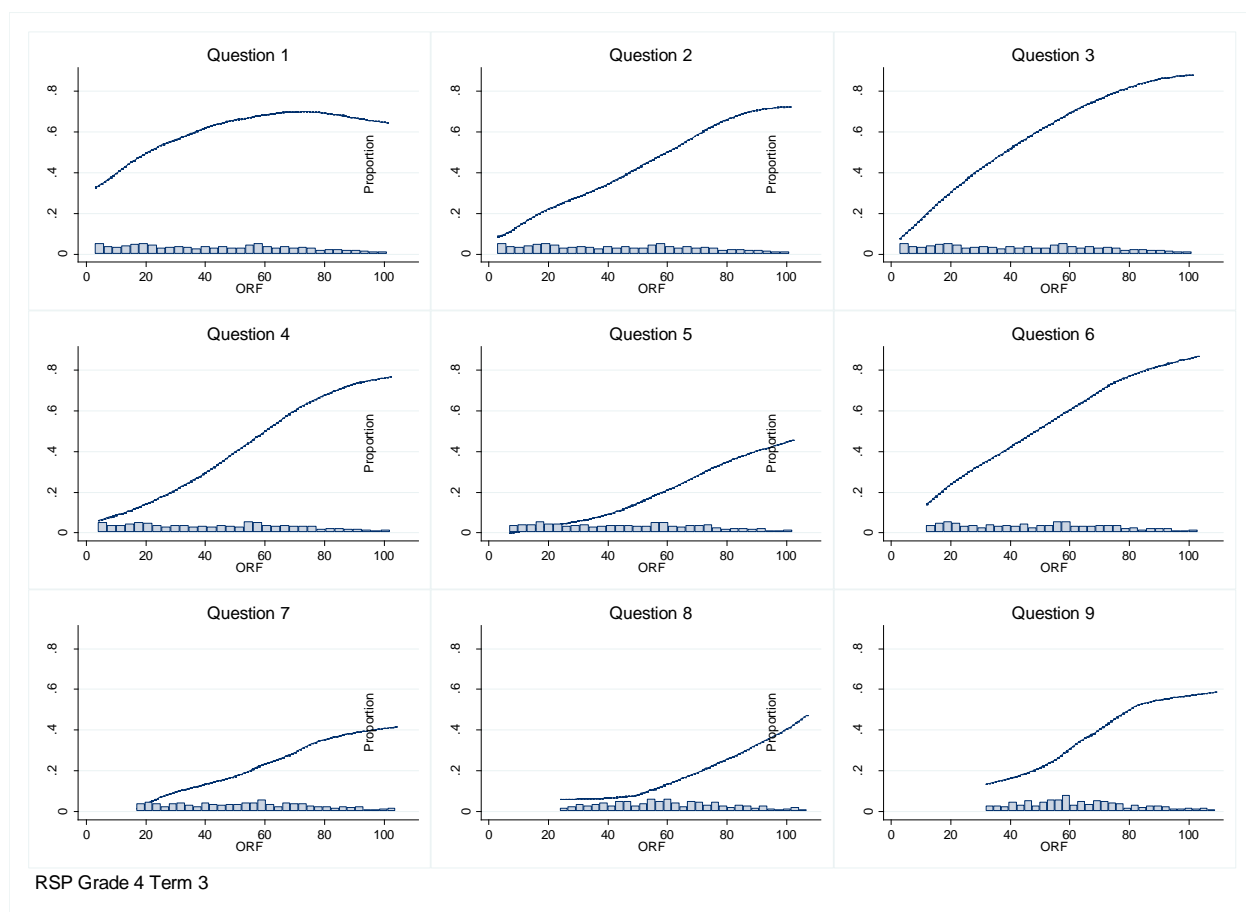
Table A 2: More details on South African EGRA-type studies used to assess reading subskills in EFAL

First Early Grade Reading Study (EGRS I) and Reading Support Project (RSP)	
Sampling of schools:	A sampling frame of 230 eligible schools developed through a process of elimination. Beginning with 458 primary schools registered in 2014 administrative data in the districts of Dr Kenneth Kaunda and Ngaka Modiri Molema we started by excluding relatively affluent schools (those in quintiles 4 and 5). Schools were excluded where the language of instruction in the Foundation Phase was not Setswana.
Sampling within schools:	20 grade 1 learners per school at baseline. Grade 1 learner cohort tracked for 5 waves (2015 to 2021). A sustainability sample of grade 3s randomly selected was added in 2018.
Stratification:	Within each stratum, randomly assigned 5 schools to each treatment group and 8 to the control group. Randomly assigned 50 schools to each treatment and 80 to the control. 10 strata of 23 similar schools based on school size, socio-economic status, and previous performance in the Annual National Assessments.
Implementer & Funder:	Implementer: Department of Basic Education. Various funders across waves including Anglo America and Zenex Foundation for waves 1-3 and USAID in waves 4-5. RSP tracks a new 2018 grade 1 cohort in a subset of the EGRS I schools. They are assessed in a second wave in 2021.
Second Early Grade Reading Study (EGRS II)	
Sampling of schools:	Representative of districts of Ehlanzeni and Gert Sibande in Mpumalanga province. Random allocation of qualifying schools to treatment and control groups. Schools are either isiZulu or Siswati LOLT schools.
Sampling within schools:	Random selection of learners from each school. +/- 20 learners per school at baseline (wave 1). Tracked for 5 waves.
Stratification:	Ten strata based on school size, performance in the Annual National Assessments (ANA) and by school wealth (DBE Quintile).
Implementer & Funder:	Implementer: Department of Basic Education. Funder: USAID
Story Powered Schools (SPS)	
Sampling of schools:	Within each district, the sampling team worked with DBE officials to create clusters of 10 schools for cohort I, and 4 schools for cohort I. Clusters were randomly assigned to treatment and control. 4 schools (cohort I)/2 schools (cohort II) were randomly selected from each cluster for inclusion in the evaluation. Schools are either isiZulu or isiXhosa LOLT.
Sampling within schools:	Random selection of one grade 2, one grade 3 and one grade 4 class per school. Random selection of 5 girls and 5 boys within each selected class
Stratification:	7 strata based on EC DoE and KZN DoE districts
Implementer & Funder:	Implementer: NORC at the University of Chicago. Funder: USAID
Funda Wande (FW) Limpopo - 2021	
Sampling of schools:	Limpopo Department of Education identified 379 quintile 1 to 3 Sepedi LOLT schools in the Capricorn North and South districts of Limpopo that did not have existing early learning outcomes interventions in grades 1 to 3, chronic management problems or learner-educator ratios above 50:1. These schools were invited to apply for the intervention and a final sample of 180 eligible schools was created. The evaluation team then randomised these 180 schools into a control group and two treatment arms.
Sampling within schools:	Random selection two grade R, grade 1, grade 3 and grade 6 classes per school. Within each class random selection of 10 grade R, 10 grade 1, 5 grade 3 and 5 grade 6 learners per class.
Stratification:	4 strata based on LP DoE districts and presence of Molteno grade R literacy programme
Implementer & Funder:	Implementer: SALDRU at the University of Cape Town. Funder: Allan Gray Orbis Foundation
Leadership for Literacy project (LFL) - 2017	

Contract Number:, Order Number:

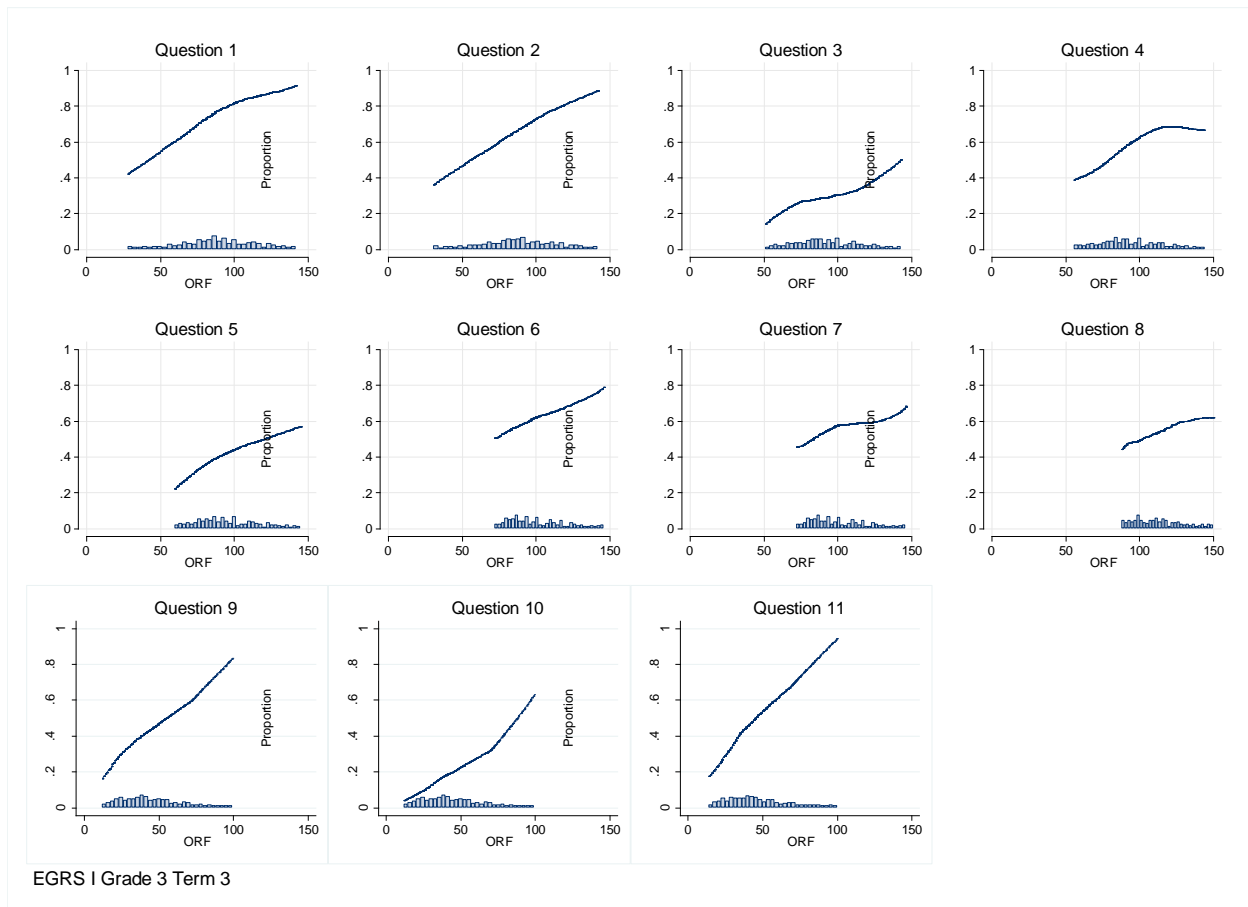
Sampling of schools:	Matched pairs design. For the isiZulu sample: 30 potentially higher performing no-fee or low fee schools (with largely isiZulu home language students) matched to under-performing pairs of schools using the Annual National Assessments. A total of 61 schools: 22 of the schools are in KwaZulu-Natal, 20 in Gauteng and 19 in Limpopo.
Sampling within schools:	Students from one grade 3 and one grade 6 class per school were sampled. Originally 10 students from each class were to be randomly sampled for the one-on-one assessments. But by wave 2 fewer or higher numbers of randomly sampled learners are achieved: 4-15 students per class in grade 6 and 6-15 per class in grade 3.
Stratification:	None
Implementer & Funder:	Implementer: Research on Socio-economic Policy (ReSEP). Funder: Economic Sciences Research Council (ESRC/DFID)

Figure A 1: Relationship between EFAL fluency and individual comprehension questions: RSP Grade 4, term 3



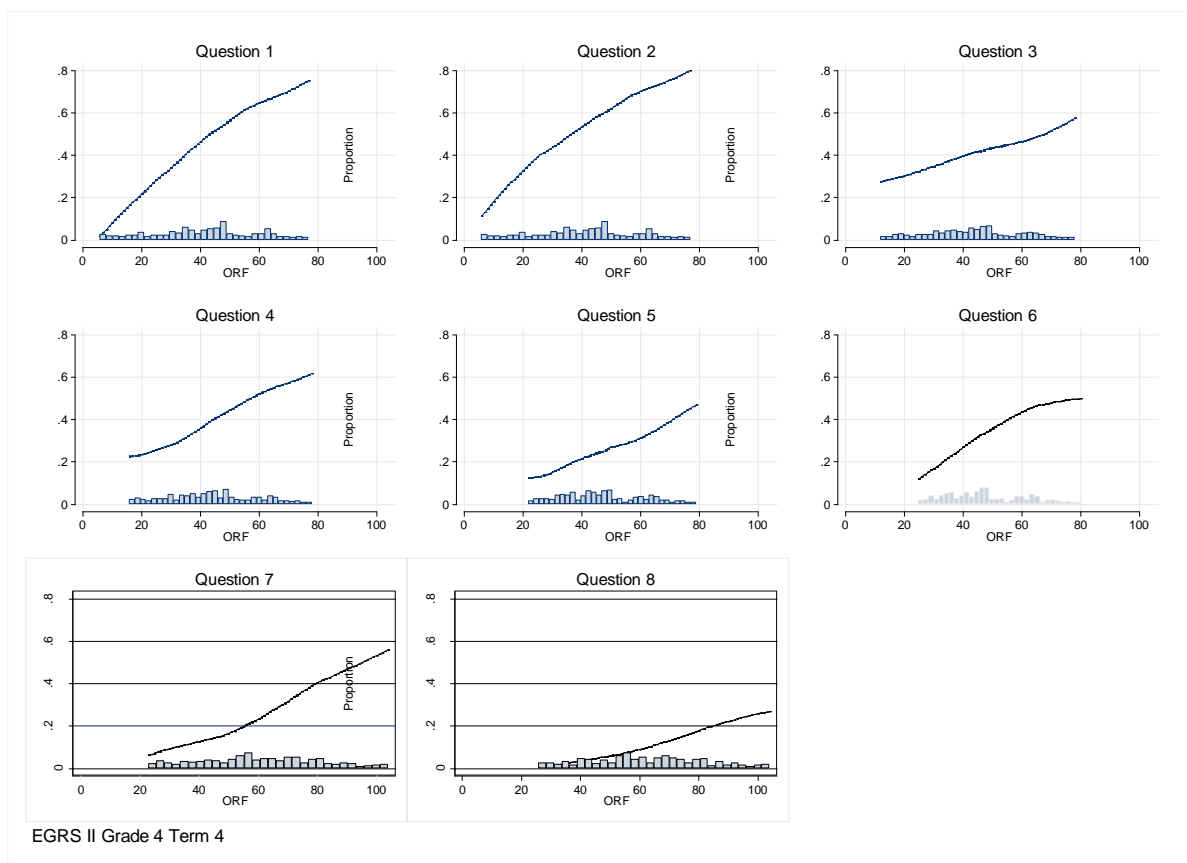
Source: RSP 2021, own calculations. Notes: The histogram bars reflect each grade sample's distribution of ORF scores. The lines are locally weighted polynomial regressions of ORF against the proportion getting the comprehension question correct.

Figure A 2: Relationship between EFAL fluency and individual comprehension questions: EGRS I Grade 3, term 3



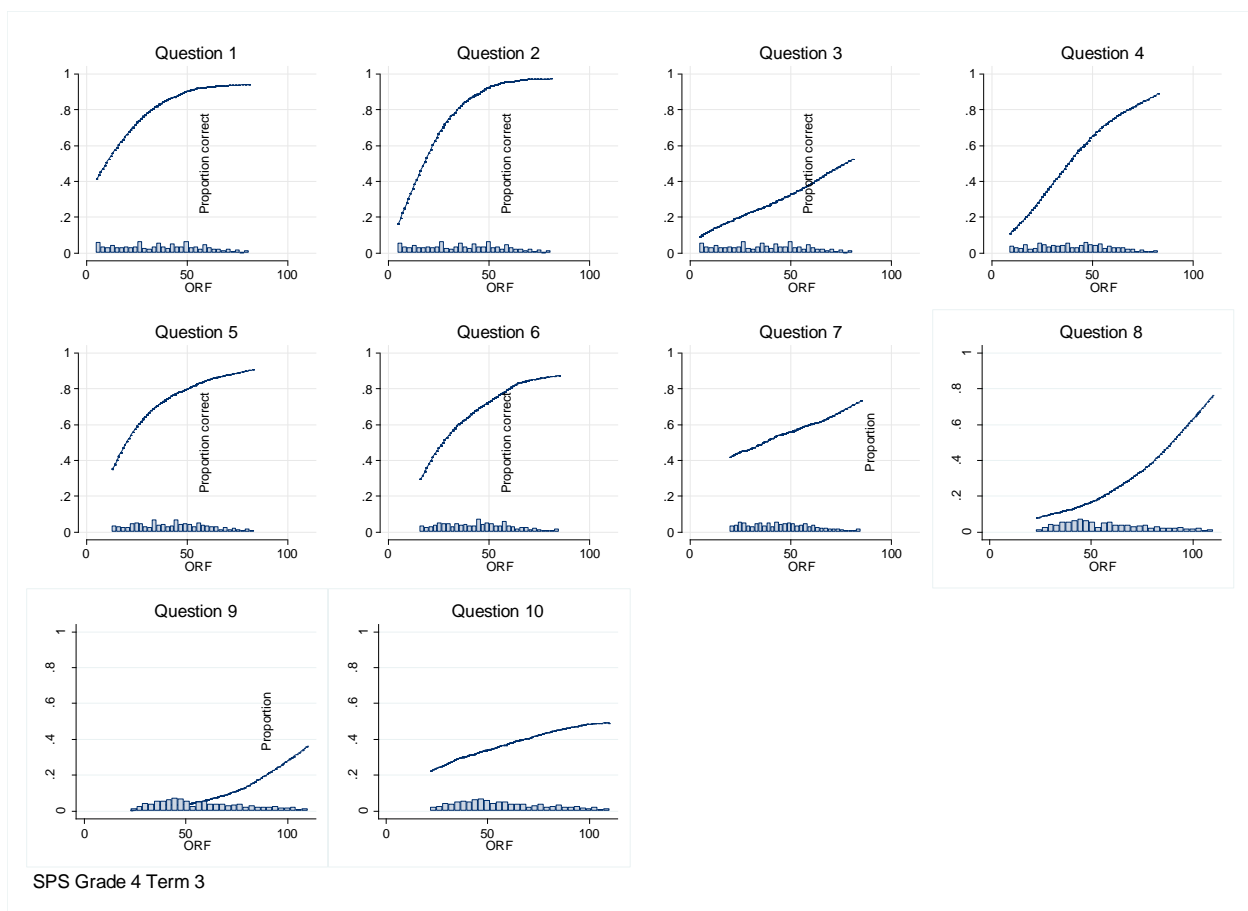
Source: EGRS I, own calculations. Notes: The histogram bars reflect each grade sample's distribution of ORF scores. The lines are locally weighted polynomial regressions of ORF against the proportion getting the comprehension question correct.

Figure A 3: Relationship between EFAL fluency and individual comprehension questions: EGRS II grade 4, term 4



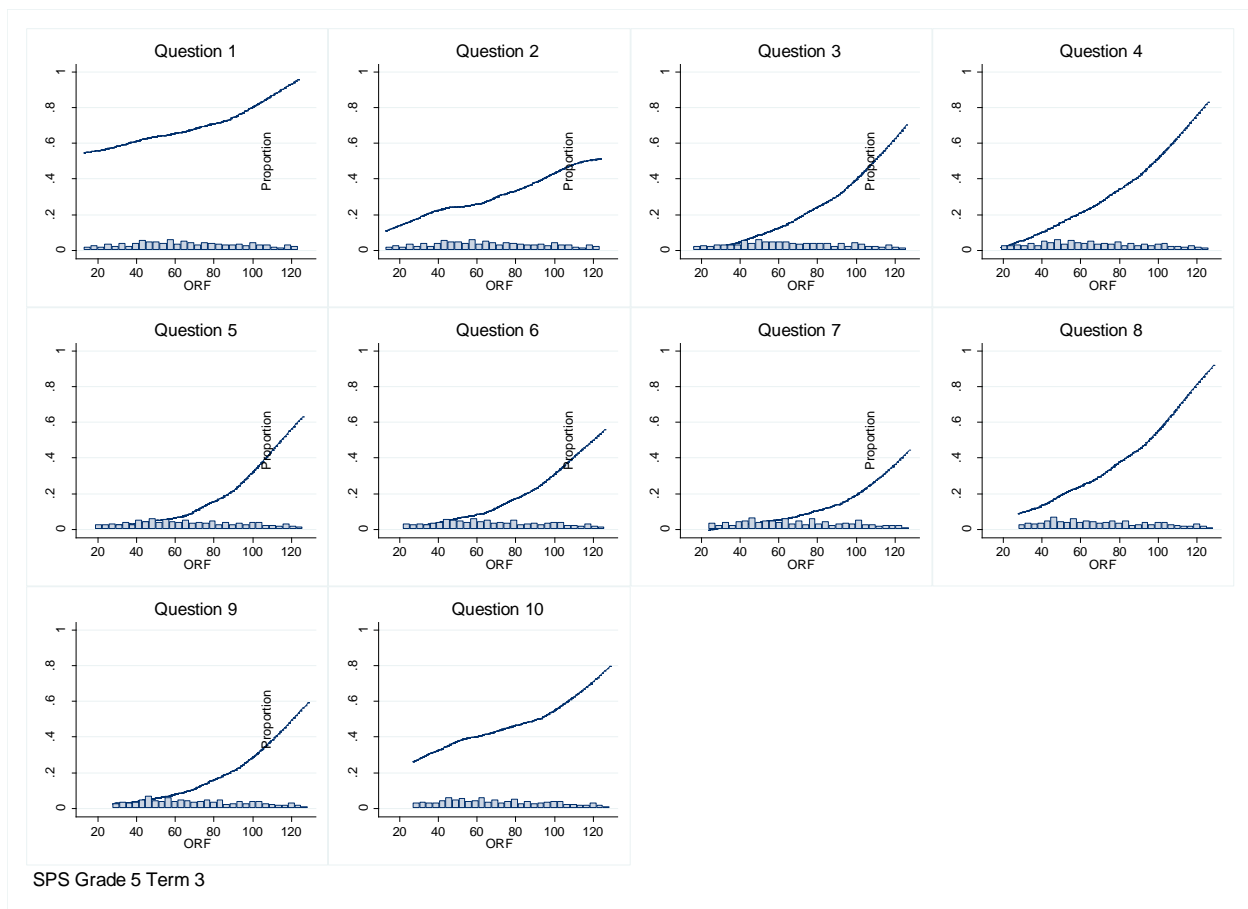
Source: EGRS II, own calculations. Notes: The histogram bars reflect each grade sample's distribution of ORF scores. The lines are locally weighted polynomial regressions of ORF against the proportion getting the comprehension question correct.

Figure A 4: Relationship between EFAL fluency and individual comprehension questions: SPS II grade 4, term 3



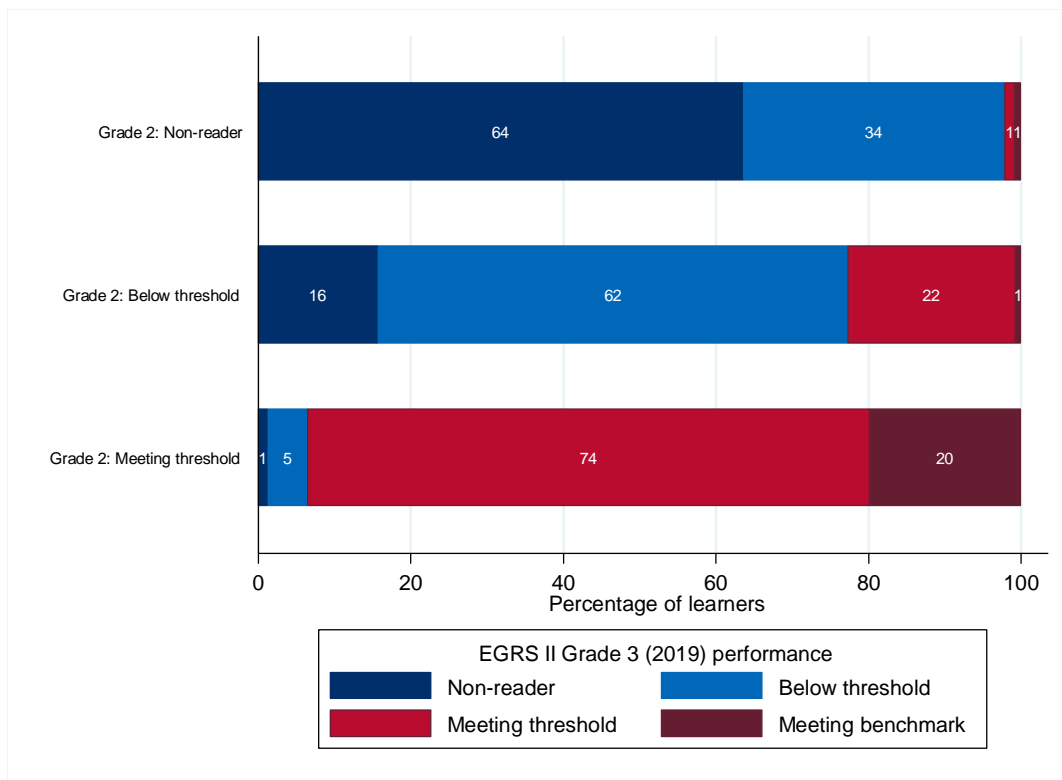
Source: SPS, own calculations. Notes: The histogram bars reflect each grade sample's distribution of ORF scores. The lines are locally weighted polynomial regressions of ORF against the proportion getting the comprehension question correct.

Figure A 5: Relationship between fluency and individual comprehension questions: SPS II grade 5, term 3, EFAL



Source: SPS, own calculations. Notes: The histogram bars reflect each grade sample's distribution of ORF scores. The lines are locally weighted polynomial regressions of ORF against the proportion getting the comprehension question correct.

Figure A 6: Fluency in grade 3, by learners' fluency profile in grade 2, EFAL



Source: EGRS II, own calculations.