Benchmarking early grade reading skills in Nguni languages

Cally Ardington, Nangamso Mtsatse & Gabrielle Wills
Joint work with Elizabeth Pretorius, Nicola Deghaye, Nompumelelo Mohohlwane, Alicia Menendez and Servaas van der Berg
Report funders

Funders of EGRA studies used as data for this report

Allan Gray Orbis Foundation Endowment

USAID

USAID

ESRC Economic & Social Research Council

UKaid

Research jointly supported by the ESRC and DFID

ZENEX F O U N D A T I O N

Leading | Connecting | Learning
A) Background
Why do we need language specific benchmarks?

Nguni languages

- Transparent
  - each letter always represents the same sound
- Large number of double and triple consonants
  - need to recognise hl, dl, kh, tsh, ndl, gcw, ntsw etc. to read early grade texts
- Agglutinating Conjunctive
  - Long words High degree of visual similarity within and between word

facilitates decoding

challenging for decoding
How do we set benchmarks?

- Which skill is benchmarked and the level are which it is set is determined by **data**
- Based on exploratory analysis of largest existing early grade reading assessment data for Nguni languages
- Makes **no assumptions** about the accuracy-speed and fluency-comprehension relationships for each language
- Sensitive to current realities of learning
- Cognisant of curriculum requirements
- Grounded in theoretical understanding of reading development
What data did we use?

• Collated 5 studies collecting early grade reading assessment data between 2017 and 2019
• Almost 16,400 unique learners in more than 660 schools
• Three Nguni languages - siSwati, isiXhosa, isiZulu

Characteristics:

Four provinces - Eastern Cape, KwaZulu-Natal, Gauteng, Mpumulanga

99% Quintile 1 to 3 schools

86% rural schools

98% of learners were tested in a Nguni language which matched the Foundation Phase LOLT in their school and their home language
Isobho Lamatshe
Kukhona isihambi esilambe kakhulu.
Sahamba sicela emizini yabantu. Abantu babengenakho ukudla.
Isihambi sathola isu. Isihambi sathola ibhodwe.
Sathatha amatshe sawafaka ebhodweni. Sathela amanzi. Sabasa umlilo, sabeka ibhodwe eziko.
Sama salinda ibhodwe laze labila.
Kwafika intombazane yacela ukwazi ukuthi siphekani isihambi eziko.
“Ngipheka isobho elimnandi lamatshe. Kodwa kumele ngilifake into ukuze linongeke,” kusho isihambi.
“Nginezaqathe mina,” wabe esenika isihambi. Sazifaka ebhodweni.

Example questions:

Yini indaba izakhamuzi zazingasini zikhumbi ukudla?
[Why did the village residents not give the traveller any food?]

Senzani isihambi gamatshe?
[What did the traveller do with the stones?]
Various taxonomies of comprehension:
We focus on PIRLS

PIRLS classifications:

i) Focus on and Retrieve Explicitly Stated Information

ii) Make Straightforward Inferences

iii) Interpret and Integrate Ideas and Information

iv) Evaluate and Critique Content and Textual Elements
Proportion answering each question correctly, examples from our EGRA data

Notice the wide range of difficulty within the literal questions.

Some literal questions are more challenging than inferential questions.

I.e. there is a wide range of difficulty within comprehension process, and no clear ordering between processes.

Similar result to PIRLS

Note: This is for learners attempting all questions
B) Establishing benchmarks
Data driven but grounded in theory

• Details in technical report
• Reading comprehension is a complex phenomenon with different processes come into play as reading proficiency increases.
  • Within each process, accuracy tends to develop first followed by speed.
  • We explicitly analyse both accuracy and speed and their interrelationship.

<table>
<thead>
<tr>
<th>Ph. Awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter-sounds</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syllable reading</th>
<th>Word reading</th>
<th>Context Fluency (ORF)</th>
<th>Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>accuracy</td>
<td>increased processing speed</td>
<td>automaticity</td>
<td>working memory free for meaning</td>
</tr>
</tbody>
</table>

Our analytical approach aligns with the decoding threshold hypothesis put forward by Wang et al. (2019)
• Until decoding occurs above a lower bound threshold level, reading comprehension is unlikely to develop/remain stagnant.
• There may also be an upper threshold, beyond which there are no additional gains in comprehension for increased decoding skills.
• Suggests that the relationship between fluency & comprehension will break down at low and high levels of fluency.
Which skills do we benchmark?

- **Letter-sound knowledge** which refers to alphabetic knowledge of the written code

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

<table>
<thead>
<tr>
<th>Accuracy*</th>
<th>Speed*</th>
<th>Prosody</th>
</tr>
</thead>
<tbody>
<tr>
<td>The percentage of words that are read correctly</td>
<td>The number of words that are attempted in a time period</td>
<td>How natural reading sounds (how it conforms to speech rhythms &amp; intonation patterns &amp; reflects punctuation conventions)</td>
</tr>
</tbody>
</table>
What are the thresholds/ benchmarks?

By the end of grade 3
all learners should be able to read
35 WORDS correct per minute
when reading a passage

By the end of grade 2
all learners should be able to read at
least 20 WORDS correct per minute
when reading a passage

By the end of grade 1
all learners should know their LETTER-SOUNDS well, sounding at least 40
correct letters per minute
C) Establishing fluency thresholds and benchmarks
How are accuracy & speed related?

Accuracy flattens out around 95%.

Low accuracy, low speed.
What does the speed-accuracy relationship look like across languages?

Across all the samples, accuracy of 95% is associated with speeds ranging from 22 to 34 words per minute.

What about slow but accurate readers?
By the end of Grade 3, between 90% & 97% of accurate readers are reading faster than 20 words per minute.
What is the relationship between fluency & comprehension? (1)
What is the relationship between fluency & comprehension? (2)

Learners attempting all questions

Fluency zone - shift to comprehension skills

Low accuracy, low fluency

Accuracy zone, Improving fluency

Oral Reading Fluency

Comprehension score
<table>
<thead>
<tr>
<th></th>
<th>isiXhosa</th>
<th>isiZulu</th>
<th>siSwati</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cannot read one word: ORF=0</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean correct letter-sounds per minute</td>
<td>12.8</td>
<td>10.8</td>
<td>18.3</td>
</tr>
<tr>
<td>% unable to sound one letter</td>
<td>12%</td>
<td>20%</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Below lower threshold: ORF=1-19 cwpm</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% with at least 95% accuracy</td>
<td>19%</td>
<td>25%</td>
<td>19%</td>
</tr>
<tr>
<td>Comprehension (% of total correct)</td>
<td>21%</td>
<td>18%</td>
<td>21%</td>
</tr>
<tr>
<td>Comprehension (% of attempted correct)</td>
<td>47%</td>
<td>51%</td>
<td>32%</td>
</tr>
<tr>
<td><strong>Meets lower threshold: ORF=20-34 cwpm</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% with at least 95% accuracy</td>
<td>71%</td>
<td>78%</td>
<td>76%</td>
</tr>
<tr>
<td>Comprehension (% of total correct)</td>
<td>46%</td>
<td>46%</td>
<td>53%</td>
</tr>
<tr>
<td>Comprehension (% of attempted correct)</td>
<td>65%</td>
<td>73%</td>
<td>62%</td>
</tr>
<tr>
<td><strong>Meets benchmark: ORF=35+ cwpm</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% with at least 95% accuracy</td>
<td>87%</td>
<td>90%</td>
<td>84%</td>
</tr>
<tr>
<td>Comprehension (% of total correct)</td>
<td>59%</td>
<td>62%</td>
<td>74%</td>
</tr>
<tr>
<td>Comprehension (% of attempted correct)</td>
<td>73%</td>
<td>78%</td>
<td>74%</td>
</tr>
</tbody>
</table>

What is the learner profile in each reading zone?

Reading classification zones provide meaningful distinctions in accuracy and comprehension across learners.

- **Low letter-sound knowledge**
- **Low accuracy, low comprehension**
- **Developing accuracy, emergent comprehension**
- **Developed accuracy, developing comprehension**
Thresholds set low enough that large enough proportions can meet these thresholds/benchmarks but are still ambitious enough to support reading development.
How do the thresholds/benchmarks relate to oral reading fluency progression?

- We have longitudinal data that allows us to follow learners as they progress through school.

- We can compare their performance at the 2nd assessment depending on whether they were meeting the reading threshold or benchmark at their 1st assessment.

- The time between the 1st and 2nd assessment ranges from 12 to 18 months.
How do the thresholds/benchmarks relate to oral reading fluency progression? (2)

The ORF thresholds and benchmark predict later fluency: Most learners who meet the threshold meet the benchmark the next time they are assessed.

### Not meeting lower threshold (Assess. 1)

<table>
<thead>
<tr>
<th>Language</th>
<th>Grade/Term</th>
<th>ORF = 1-19 (Assessment I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>isiXhosa</td>
<td>Grade 2 Term 1</td>
<td>8 26 50 17</td>
</tr>
<tr>
<td></td>
<td>Grade 2 Term 1</td>
<td>3 38 51 7</td>
</tr>
<tr>
<td></td>
<td>Grade 3 Term 1</td>
<td>3 29 55 12</td>
</tr>
<tr>
<td></td>
<td>Grade 4 Term 1</td>
<td>4 29 63 4</td>
</tr>
<tr>
<td>isiZulu</td>
<td>Grade 2 Term 1</td>
<td>5 21 53 21</td>
</tr>
<tr>
<td></td>
<td>Grade 2 Term 4</td>
<td>2 30 55 13</td>
</tr>
<tr>
<td></td>
<td>Grade 3 Term 1</td>
<td>5 25 47 23</td>
</tr>
<tr>
<td></td>
<td>Grade 3 Term 1</td>
<td>5 27 57 15</td>
</tr>
<tr>
<td></td>
<td>Grade 4 Term 1</td>
<td>4 30 50 16</td>
</tr>
<tr>
<td>siSwati</td>
<td>Grade 2 Term 4</td>
<td>5 33 50 12</td>
</tr>
</tbody>
</table>

### Meet lower threshold (Assess. 1)

<table>
<thead>
<tr>
<th>Language</th>
<th>Grade/Term</th>
<th>ORF = 20-34 (Assessment I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>isiXhosa</td>
<td>Grade 2 Term 1</td>
<td>2 32 66</td>
</tr>
<tr>
<td></td>
<td>Grade 2 Term 1</td>
<td>4 39 57</td>
</tr>
<tr>
<td></td>
<td>Grade 3 Term 1</td>
<td>2 35 62</td>
</tr>
<tr>
<td></td>
<td>Grade 4 Term 1</td>
<td>4 47 51</td>
</tr>
<tr>
<td>isiZulu</td>
<td>Grade 2 Term 1</td>
<td>4 23 72</td>
</tr>
<tr>
<td></td>
<td>Grade 2 Term 4</td>
<td>2 31 59</td>
</tr>
<tr>
<td></td>
<td>Grade 3 Term 1</td>
<td>5 38 78</td>
</tr>
<tr>
<td></td>
<td>Grade 3 Term 1</td>
<td>2 33 65</td>
</tr>
<tr>
<td></td>
<td>Grade 4 Term 1</td>
<td>4 38 60</td>
</tr>
<tr>
<td>siSwati</td>
<td>Grade 2 Term 4</td>
<td>3 45 51</td>
</tr>
</tbody>
</table>

### Oral Reading Fluency (Assessment II)

- **0**
- **1-19**
- **20-34**
- **34+**
Learners not meeting the lower threshold (cwpm < 20) by the beginning of grade 4 have very poor written comprehension skills in grade 5.

Learners achieving at least 5 out of 8 questions correct in Grade 5 were typically reading above 20 words per minute at the beginning of Grade 4.

How do the thresholds/benchmarks relate to written comprehension?

Previous ORF distribution by current written comprehension score

<table>
<thead>
<tr>
<th>80</th>
<th>70</th>
<th>60</th>
<th>50</th>
<th>40</th>
<th>30</th>
<th>20</th>
<th>10</th>
<th>0</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>isiXhosa</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>isiZulu</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

Shows distribution of ORF score at the beginning of grade 4, by grade 5 term 3 written comprehension.
D) Establishing a letter-sound benchmark
Why should letter-sounds be benchmarked?

Letter-sounds predict future oral reading fluency

[Graph showing the relationship between letter sounds and oral reading fluency]
At what level should we set the letter-sound benchmark?

Accuracy and speed increase steadily and then accuracy flattens out.
At what level should we set the letter-sound benchmark?

There are diminishing improvements in letter-sounds.
Why are double and triple consonants important?

We don’t provide benchmarks for double or triple consonants, but knowledge of them is vital to be able to read any Foundation Phase text. Currently very few learners at the end of grade 1 can sound these. Double and triple consonants should be taught and assessed as a distinct task.
E) Q&A panel
Why?

- Good early predictor of oral reading fluency (ORF) later in Foundation Phase.
- Improvements in letter-sound speed stagnate around this point.
- Needs to be low enough to measure incremental progress.
- Needs to be ambitious enough to support curriculum demands and improved reading outcomes.

What does the benchmark look like in practice?
• **Instructional focus?**
  • learners *not* reaching benchmark: letter identification and phoneme-grapheme recognition
  • learners meeting benchmark: word identification to improve decoding skills

• **Who is currently meeting this benchmark?**
  • By the beginning of grade 2, between 10% and 45% of learners in this sample had reached this benchmark
Why?

- Below this threshold, accuracy is poor & we find little evidence that learners can comprehend what they have read.
- This is a minimum threshold. If learners do not reach this level of fluency, higher order reading skills are very unlikely to develop.

What does the threshold look like in practice?

Instructional focus for learners meeting threshold

- practice with text to recognize words more quickly, as well as to improve their comprehension

Who is currently reaching this threshold?

- By the end of grade 3, between 53% and 76% of the learners in this sample had reached this grade 2 threshold.
Why?
• At this level of fluency reading comprehension becomes increasingly possible when learners read on their own.
• Once learners reach this level of fluency, it appears that poor comprehension skills become the limiting factor to further literacy development.

What does this benchmark look like in practice?

**SISWATI**

**ISIZULU**

**ISIXHOSA**
Instructional focus for learners meeting this benchmark:
- skills and strategies to improve their understanding of and engagement with the text
- encouraging vocabulary development to support comprehension
- fluency skills should continue to improve from this milestone.

Who is currently meeting this benchmark?
- By the end of grade 3, approximately only a quarter of learners had reached the benchmark.
Further questions?

**Foundation Phase**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Non-reader</th>
<th>Emerging reader</th>
<th>Developing proficiency reader</th>
<th>Competent reader</th>
<th>Skilled reader</th>
</tr>
</thead>
</table>

**Intermediate Phase**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Grade R</th>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 5</th>
</tr>
</thead>
</table>

- **By the end of grade 1**
  - All learners should know their LETTER-SOUNDS well, sounding at least 40 correct letters per minute.

- **By the end of grade 2**
  - All learners should be able to read at least 20 WORDS correct per minute when reading a passage.

- **By the end of grade 3**
  - All learners should be able to read 35 WORDS correct per minute when reading a passage.
Additional slides
### How comparable were the different passages?

<table>
<thead>
<tr>
<th>Story</th>
<th>Term</th>
<th>Grade</th>
<th>Language</th>
<th>Number of syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>siSwati</td>
<td>31</td>
</tr>
<tr>
<td>B</td>
<td>IV</td>
<td></td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>B</td>
<td>IV</td>
<td>3</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>C</td>
<td>IV</td>
<td>1</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>A</td>
<td>I</td>
<td>2</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>A</td>
<td>IV</td>
<td>3</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>C</td>
<td>I</td>
<td>3</td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>B</td>
<td>I</td>
<td>4</td>
<td>isiXhosa</td>
<td>32</td>
</tr>
<tr>
<td>A</td>
<td>III</td>
<td></td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>B</td>
<td>III</td>
<td>5</td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>B</td>
<td>I</td>
<td>2</td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>B</td>
<td>IV</td>
<td>3</td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>B</td>
<td>III</td>
<td>4</td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>B</td>
<td>IV</td>
<td>5</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>B</td>
<td>III</td>
<td>6</td>
<td>isiZulu</td>
<td>35</td>
</tr>
<tr>
<td>D</td>
<td>IV</td>
<td>4</td>
<td></td>
<td>28</td>
</tr>
</tbody>
</table>

Notes:
- A, B, C, D reflect which story was used
- Project/Study: EGRSII, SPS, FW, LFL
- Term reflected by I, II, IV
- Language reflected by S = Siswati, X = isiXhosa, Z = isiZulu
What do we mean by comprehension?

CAPS Foundation Phase Home Language refers to the following comprehension levels:

- Literal
- Reorganisation
- Inferential
- Evaluation
- Appreciation
What about slow but accurate readers?

By the end of Grade 3, between 90% and 97% of accurate readers are reading faster than 20 words per minute.
How do the thresholds/benchmarks relate to oral reading fluency progression?

Start with learners who could not read one word at the first assessment.

What do they look like when we see them 12 – 18 months later?
How do the thresholds/benchmarks relate to written comprehension?

Learners achieving at least 5 out of 6 questions correct in Grade 3 were typically reading above 20 words per minute at the end of Grade 2.

Learners not meeting the lower threshold by the end of grade 2 have very poor written comprehension skills in grade 3.

Learners achieving at least 5 out of 6 questions correct in Grade 3 were typically reading above 20 words per minute at the end of Grade 2.

Shows distribution of ORF score at the end of grade 2, by grade 3 term 4 written comprehension.
How does the letter sounds benchmark relate to future oral reading fluency?

Less than 20 letter sounds per minute (Assess. 1)

20-39 letter sounds per minute (Assess. 1)

Meets benchmark of 40+ (Assess. 1)
Who is meeting the letter-sounds benchmark? (1)

### isiZulu correct letter sounds per minute

<table>
<thead>
<tr>
<th></th>
<th>SPS</th>
<th>EGRS</th>
<th>LFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 2 Term 1</td>
<td>4.6</td>
<td>16.3</td>
<td>35.0</td>
</tr>
<tr>
<td>Grade 3 Term 1</td>
<td>8.2</td>
<td>14.0</td>
<td>23.8</td>
</tr>
<tr>
<td>Grade 3 Term 2</td>
<td>8.3</td>
<td>30.8</td>
<td>22.6</td>
</tr>
<tr>
<td>Grade 3 Term 3</td>
<td>13.4</td>
<td>53.5</td>
<td>17.0</td>
</tr>
<tr>
<td>Grade 4 Term 1</td>
<td>12.4</td>
<td>5.5</td>
<td>15.8</td>
</tr>
<tr>
<td>Grade 4 Term 2</td>
<td>8.8</td>
<td>15.0</td>
<td>24.0</td>
</tr>
<tr>
<td>Grade 4 Term 3</td>
<td>9.6</td>
<td>25.0</td>
<td>27.0</td>
</tr>
<tr>
<td>Grade 4 Term 4</td>
<td>9.1</td>
<td>15.0</td>
<td>29.0</td>
</tr>
</tbody>
</table>

**Large variation across studies**

Correct letter sounds per minute

- **0**
- **1-19**
- **20-29**
- **30-39**
- **40+**
Who is meeting the letter-sounds benchmark? (2)
Do learners know their complex consonant sequences?