unesco
Institute for Statistics

25 YEARS

Education Data and Statistics Commission

WG/GAML/11/2.1

TECHNICAL DOCUMENTATION SUPPORTING EXEMPLAR EARLY GRADES READING ITEMS

DRAFT

Global Alliance to Monitor Learning

1. Background [1]

During the benchmarking meeting of the Technical Advisory Group (TAG) for SDG 4.1.1.a in May 2024, a recommendation to calibrate assessments to benchmarks meaningfully to the MPLs and to each other was proposed, more specificity regarding the degree of item difficulty used in newer assessments such as EGRA, FLM, and/or the assessments associated with the PAL network. AMPL-a has already done this with pairwise linking. Newer assessments currently link (or have proposed to link) items to the MPLs through a benchmark such as the percent of correctly answered comprehension questions or the percent of correctly answered numeracy questions. But without delimiting the level of difficulty of the items or their predictive ability, these benchmarks are not as meaningful or comparable as they could be.

Another important reason to better calibrate the difficulty level across assessments is to increase the utility for country stakeholders. Well calibrated assessment items can mitigate a country's discouragement by poor results or, conversely, complacency if the items are too easy.

Finally, one ought to expect the newer assessments to yield results that are broadly comparable, in terms of the percentages said to be at the minimum level or above, to more known assessments used in LICs, namely PASEC and ERCE, and a new one that has been proposed and used in a few LICs, namely AMPL-a, but that falls more in the family of traditional assessments. That said, the TAG also noted that one should avoid the temptation to over-specify, as one is not creating a new global assessment.

2. Overview

Learning to read is a multifaceted process that requires the simultaneous development of various skills that lead to proficient reading. Proficient readers can read fluently and understand the meaning of what they read. To reach proficiency students must have adequate oral language and background knowledge and must have learned to read words- including multisyllabic words- with automaticity. This happens when words have been orthographically mapped and are stored in long-term memory. The cognitive processes needed to learn to read are the same regardless of language, at least for non-ideographic languages. However, because reading is a language-based process, some specific precursor skills needed to learn to read can differ between languages, and as a result the specific assessment tasks that best predict reading development. For example, in syllabic languages syllable reading predicts later reading better than letter sounds. Despite these differences, assessments of precursor skills and higher order skills can be used to determine students' progress toward reading proficiency. Four reading components are addressed throughout

[1] Authored by Sylvia Linan-Thompson under the guidance of Luis Crouch.

this report, grapheme/syllable sound correspondence, word reading, reading accuracy and fluency and reading comprehension.

2.1 Reading Components

There is general agreement that learning to read requires the development of several skills that range from producing letter or syllable sounds to building word reading automaticity in isolation and in connected text to developing reading comprehension skills and strategies needed to read proficiently. This section has two parts. The first provides a brief definition of the reading components that are included in this review and the second describes assessment tasks typically used to assess them.

2.1.1 Letter Sounds

A critical first step in learning to read is to recognize, name, produce the sound, and write the letters or symbols of the alphabet (McBride-Chang, 1999; Puranik et al., 2011, Scanlon et al., 2010). While assessing these skills can provide information about students' evolving abilities, knowledge of letter sounds is most predictive of the ability to sound out words and future reading success. Students who can map sounds to letters learn that language is made up of discrete sounds, which is true regardless of the orthographic structure of the language (i.e. Bajre & Khan, 2019). Once students have a few established symbol/ sound correspondences, they can begin to decode words.

2.1.2 Word Reading

Automatic word reading or effortless word reading is essential to reading comprehension. The ability to instantly recognize words-both their pronunciation and meaning allows students to devote cognitive resources to comprehend what they are reading rather than devoting scarce cognitive resources to effortful decoding of symbols. This relationship is particularly important when students are learning to read (Vellutino, Tunmer, Jaccard, & Chen, 2007). This relationship weakens as students become proficient readers (Hoover & Gough, 1990; Joshi, Williams, & Wood, 1998; Rupley, Willson, & Nichols, 1998).

2.1.3 Text Reading Accuracy and Fluency

Automatic word reading should transfer from reading words in isolation to reading them with equal automaticity in connected text: fluent reading. The generally accepted definition of fluent reading is 1) accurate reading at 2) an appropriate rate and 3) with expression. For fluent reading and to gain meaning from text, all three of these components must be in place (Kuhn & Stahl, 2000).

Reading accuracy, the ability to read words correctly, is sometimes used to gauge student progress toward proficiency. Reading accuracy is necessary to understand what is read. If there are too many errors, a student may build an inaccurate or incomplete representation of what is read. Therefore, to comprehend, a student must be able to read 95% of the text accurately (Hasbrouck & Glaser, 2019; Treptow, Burns, & McComas, 2019). This is the level at which a student can learn, and 98% accuracy is the level at which a student can read

independently. There is agreement that it is not enough to be accurate; reading skills must be automatic.

2.1.4 Listening Comprehension

Broadly speaking, listening comprehension refers to the ability to understand, interpret, and construct meaning from spoken language. With regard to literacy, the focus is on the comprehension of auditory text or texts read to the student. Listening comprehension is a complex cognitive process, that requires that a student recognize words understand the semantic and syntactic aspects of sentence structure, and grasp the overall message being conveyed. To accomplish this, students must be focused and engaged with the information presented. Through listening comprehension tasks, students can develop comprehension skills and strategies such as identifying the main idea, drawing conclusion and making inferences. The process is facilitated when students listen to text in the language they speak best. Listening comprehension predicts both early and later growth of reading comprehension. Students who are not yet proficient readers can benefit from opportunities to develop listening comprehension.

2.1.5 Reading Comprehension

Reading comprehension is identified as the goal of reading. When a student can comprehend what is read, they simultaneously extract and construct meaning through their interaction with text (Snow, 2002). Students learning to read develop and practice foundational skills until reading comprehension becomes both an automatic, unconscious process and a controlled strategic process; but how well they comprehend depends on a match between reader and text (Kintsch, 2004). While explicitly taught comprehension skills and strategies are critical, a student's breadth and depth of vocabulary knowledge and their background knowledge also impact reading comprehension. As noted previously, students must not only decode words to comprehend text, but they must also know the meaning of the words they read. Adequate reading comprehension depends on a person already knowing the meaning of 90-95% of the words in a text (Nagy & Scott, 2000). Having a threshold of knowledge about a topic helps students construct a meaningful mental model of what the text is about. This is critical for comprehension (Smith, Snow, Serry, & Hammond, 2021). A high degree of background knowledge helps weaker readers compensate for their relatively weak reading skills (Smith, Snow, Serry, & Hammond, 2021), indicating that providing only texts on topics known to students may lead to overestimates of their reading ability.

2.2 Assessment of Reading Components

Like reading development, the assessment of reading is a complex task that is impacted by language. Therefore, one set of benchmarks or a single set of text characteristics may not be appropriate for all languages. This section describes assessments typically used to assess reading development in existing exemplar assessments. Appendix A describes GPF reading constructs that can be used for SDG 1.1a reporting.

2.2.1 Letter Sounds

Letter sound tasks assess students' understanding of the connection between letters or symbols and their corresponding sounds. The purpose of assessing students' letter sound knowledge is to not only determine how well they are developing this knowledge but also provides information on whether they know enough letter sounds to read words.

Letter sound knowledge is typically assessed the first year a student enters school by presenting a student an array of letters/symbols. Students are asked to produce the sound of each letter in succession. In languages with deep orthographies in which letters many have multiple sounds or sounds are represented with various spellings, syllabaries with a large number of symbols, or those with large alphabets often extend the assessment of letter sound knowledge into a second year. In this case, assessments developed for the first year of school, include the most common letters or symbols; those typically taught in that year. The remaining symbols are assessed in the second year or beyond. Other factors to consider are the order of the letters and the number of letters included. The order of letters can be random or from most common/easiest/frequent to less common/harder/less frequent. Presenting letters in alphabetical order is not recommended, as it may lead to students simply repeating aloud the memorized alphabet, which is not reading. The number of letters may also differ. While the array used typically has 50 or 100 letters or symbols some assessment may include fewer letters or symbols. When the number of letters/symbols in the array is greater than the number of letters/symbols in the alphabet, some letters are repeated.

Most assessments of letter sound knowledge are timed. Students are asked to produce the sounds of an array letters in one minute. Each correctly produced sound is counted and the correct number of sounds produced in one minute is the score. Untimed measures are typically criterion measures. Students are asked to produce the sounds of specific letters that students of that grade level are expected to know.

2.2.2 Word Reading

Reading words in isolation measures a student's ability to quickly and accurately recognize words without needing to sound them out. This is often referred to as single word reading ability. The purpose of assessing students' word reading ability is to determine if students can read common words with automaticity.

Single word reading is typically assessed during students' first three years of school. Students are presented an array of 50 to 100 common, grade appropriate words. Even when selecting common, grade appropriate words, the length of the words used may vary based on language differences. The ordering of the words is another consideration. Like the letters in the letter sound task, words can be ordered by length (number of letters or syllables), frequency of use, or randomly (but within a limit of frequency of use in the grade being assessed). Single word reading is sometimes assessed using non-words to ensure that students use decoding skills to read the words. Both nonword reading and word reading tasks can predict later text reading and reading comprehension.

Most assessments of students' single word reading ability are timed. Students are asked to read words for one minute. Each word read correctly is counted and the correct number of words read correctly in one minute is the score. Another approach is to say a word and ask the student to identify the word from a list of words in a booklet.

2.2.3 Reading Fluency and Accuracy

Oral reading fluency tasks measure students' ability to read a connect text accurately and at a good rate. The assessment of oral reading fluency usually begins once students can read common words found in grade level texts with automaticity: this is generally between the first and second year of formal schooling. Reading fluency continues to develop through primary school (Hasbrouck & Tindal, 1992; Horn & Manis, 1987). Reading accuracy is a component of reading fluency. However, reading accuracy can be assessed independent of rate of reading. Reading accuracy tasks evaluate how accurately a student can read a text and decode words with or without automaticity. It can also provide insight into students' use of decoding strategies if these are noted during the assessment. This task may be more appropriate for students who are still developing automatic word reading skills.

Assessments of oral reading fluency are timed. Students are asked to read a 50–75-word text for one minute. Each word read correctly is counted and the correct number of words read correctly in one minute is the score. In practice most assessments of reading in recent years have given the children more than one minute to read—some assessments leave 3 minutes or so for a passage that a fluent child should read in one minute, others are even more lenient. But a common practice is to put a mark at the one-minute point so as to enable a calculation of fluency. Others do not do that either and simply measure accuracy. Assessments of reading accuracy are usually not timed. Students are asked to read a 40 - 50-word text. The score is the percentage of words read correctly out of the total words read. The length of the text will vary based on the word and sentence length common in languages with different features.

2.2.4 Listening Comprehension

Listening comprehension measures a student's ability to understand spoken language. Reading and listening comprehension draw on the same language processes and are often similar when decoding skills are controlled for (Hogan et al., 2014). However, because of the fleeting nature of auditory text, the length of the text is an important factor. Texts used for listing comprehension assessments are usually shorter than the texts used for reading comprehension.

Listening comprehension assessments usually consist of a short (4-6 sentence) text being read to or played from a recording device. Students are then asked questions one at a time. Questions can be either retrieval questions (who what when, or where) or questions (how or why) that require interpretation of the text. The score is the number of questions

answered correctly. The length of the text can vary based on the word and sentence length common in languages with different features.

2.2.5 Reading Comprehension

The purpose of a reading comprehension assessment is to measure how well a student understands a text that they have read. The ability to comprehend text develops early when provided the opportunity to engage with text and becomes more strategic as students become more proficient. Reading comprehension can be assessed at the sentence or text level.

Similar to the assessment of reading fluency and accuracy, and listening comprehension, assessing reading comprehension is impacted by the text used as a prompt. Text characteristics such as word length or frequency, text cohesion and sentence length, as well as language structure, language conventions and clarity, and knowledge demands on the reader (Davidson, 2013) contribute to text complexity. This is true whether using a text as the basis to answer questions or for a maze task. Comparability of text prompts is often defined by the length of the text and by type of sentences used as well as grade-level comparisons.

In addition, when determining the difficulty of the response stimulus, question difficulty must be considered. The most common task used to measure reading comprehension is to ask a student to read a grade level text either silently or aloud and then answer questions about the text. Questions can range from literal to integrative. However, in early primary school grades, questions are typically literal, responses to who, what, when, or where questions, that require students to retrieve information that is explicitly stated in the text. As a student progresses through primary school, inferential and evaluative questions become more common.

Comprehension at the sentence level measures comprehension of a sentence. The student reads a set of short, simple unrelated sentences silently then responds by either verifying if the sentence states a true statement or not or by answering a single question about the content. Sentence comprehension is usually not timed but the task usually takes 3 minutes. Sentence level tasks may be more appropriate for beginning readers who may not have developed the foundational skills needed to read a passage. The score is the number of correct responses.

Reading comprehension of text assessments can be either timed or untimed. For timed measures, students are asked to read a 50–75-word text for one minute. After one minute, students are asked to answer 5 questions aligned to the amount of text that was read, so that questions that are beyond where the student read are not asked (and are then either left as missing or marked wrong—there are important implications for this). For untimed measures, students are given between 3-5 minutes to read a text. They are then asked to answer 5 questions about the text. The length of the text can vary based on the word and sentence length common in languages with different features. Note that the current

guidance from UIS requires that there be a minimum of 10 comprehension items (most likely from two passages), if one is to use the assessment for global reporting. Other UIS documentation deals with this issue.

Multiple choice is another common method of assessing reading comprehension. Maze is a multiple-choice test. As students read a text with every 7th word missing, they are asked to choose the correct word from a set of three words. A student's score is the number of correctly identified words. Traditional multiple-choice assessments are also used. In these assessments students read a short text silently then answer multiple choice questions about the text. The types of questions used in these assessments are similar to those used in other assessments.

3. Item Difficulty

AIR provided item difficulty data for 33 different assessments however, only three EGRA booklets with matching data were provided limiting conclusions. Assessment data was used to determine item difficulty for reading comprehension and foundational skills (letter sounds/ syllable sounds, word reading, oral reading fluency and accuracy, and listening comprehension) that lead to reading proficiency. Of these, letter sounds/ syllable sounds, word reading, oral reading accuracy, and listening comprehension are aligned with the Global Proficiency Framework. AIR also provided benchmarks by language groups.

3.1 Range of difficulty

Guidance for developing assessment tasks used to evaluate student progress toward reading proficiency is available from organizations that have developed assessments. Although there is overlap for some tasks, others vary. The GPF also provides general guidance on item construction to reduce variability in item difficulty. However, it is important to note that student s included in this report varied on a number of characteristics beside language including content of their literacy instruction and their literacy knowledge and skills. Other variables that impact student performance on assessments are school variables (i.e. teacher training or availability of materials) and community variables (availability of environmental print or socio-economic levels). Since p-values are sample dependent, differences in p-values are the result of differences in the skills of the sample who responded to the items. In turn, student skills are impacted by school and community variables.

3.1.1 Foundational Skills

Tasks such as letter/sound or symbol/sound correspondence are easier to calibrate because there is a finite set of items and difficulty, usually defined as frequency in the language or in grade level words, should be easy to ascertain. However, data indicate that letter sound tasks range from difficult to moderate in Arabic but are difficult in Bantu languages. Syllable reading in Bantu ranges from difficult to moderate, while in languages spoken in India ranged from moderate to easy. Even though items are finite, differences in item difficulty can be due to differences in student exposure to print, lack of instruction,

and opportunities to practice. For example, in some countries letter names but not sounds are taught. Yet, it is the GPF construct, D1.1 Identify symbol/sound/fingerspelling and/or symbol/morpheme correspondences, that is acceptable for reporting for SDG 4.1.1a – reading.

Table 1. Range of symbol/sound correspondence difficulty across language groups

Language	Arabic	Bantu	Dravidian	Indo-Aryan	Sino-Tibetan
Correct Letter Sounds per Minute	.13 - .46	.09 - .13			
Correct Syllable Sounds per Minute		.04 - .32	.68 - .73	.68 - .90	.78 - .97

Likewise, the difficulty of words used for word reading tasks is relatively easy to establish. Words used in these tasks are commonly found in grade level texts or are used by students at a specific grade level. This is true whether students read the words from an array of words or identify them from a list of words after being provided the word orally. Words used for nonword tasks should follow the orthographic rules of the language, and the construction in terms of length and letters/syllable used should be similar to those used for real words.

Table 2. Range of word reading difficulty across language groups

Language	Arabic	Bantu	Dravidian	Indo-Aryan	Sino-Tibetan
Correct Word Reading per Minute		.18 - .32	.62 - .64	.61 - .86	.57 - .82
Correct Non-word Reading per Minute	.07 - .23	.04 - .24	.52 - .72	.54 - .80	.48 – .80

In languages spoken in India, word reading and non-word reading tasks ranged in difficulty from moderate to easy, indicating that either would be appropriate. Of the two tasks, GPF construct, D1.2 decode isolated words, only word reading is acceptable for reporting for SDG 4.1.1a – reading.

3.1.2 Higher Level Skills

For assessments of higher-level skills, the prompt is a text—either a sentence or a passage. As a result, attention to the construction of the prompt is critical. Passages used for reading fluency and accuracy are typically 50-70 words, but length can vary be language. Most passages are narrative or descriptive. Narrative passages have at least one character involved in a dilemma that is resolved. For narrative passages, the story structure should follow that common in the culture but should be limited to no more than three characters. Although descriptive passages, those that describe a person, place, or thing, are used, they are harder that narrative passages because they lack a clear story structure.

3.1.2.1 Oral Reading Fluency and Accuracy

Though few assessments administered both fluency and accuracy measures, available data indicates that reading accuracy assessments fall in the moderate to easy range while oral reading fluency are in the difficult to moderate range. The differences between language groups may be due to differences in the complexity of the languages, student ability to read, or student access to a variety of texts. Although, GPF construct, D2.1, say or sign a grade level continuous text at pace and with accuracy, includes both accuracy and rate, only accuracy is included in the subconstruct for second grade.

Table 3. Range of text reading difficulty across language groups

Language	Arabic	Bantu	Dravidian	Indo-Aryan	Sino-Tibetan
Oral Reading Fluency	.09 -.43	.03 - .29	.35 - .51	.34 -.76	.37 - .72
Oral Reading Accuracy	.54 - .81				

3.1.2.2 Comprehension

The questions are the response stimulus for both listening and reading comprehension tasks. Generally, literal questions (who, what, when, and where), are easier than interpretive questions (how and why) and inferential questions regardless of the response format (open-ended or multiple-choice). Current assessments ask a range of questions beginning in first grade. Across 33 different assessments, questions ranged from easy to difficult. The number of questions in each range is included in Table 4. Listening comprehension question are combined because the actual questions were not available so the type of question cannot be determined. The same is true for the reading comprehension questions.

Table 4. Distribution of questions by level of difficulty

	Low p-values (most difficult)			Mid-range p-values (medium difficult)			High p-values (least difficult)		
	.1-.19	.2-.29	.3-.39	.4-.49	.5-.59	.6-.69	.7-.79	.8-.89	.9-1
C2.1 and C3.1	1	1	13	4	6	9	6	9	4
R1.2	6	9	15	24	56	64	62	30	4

Across all the data available question difficulty ranged from .19 to 1 for reading comprehension and .20 to .99 for listening comprehension.

Table 5. Range of comprehension question difficulty across language groups

Language	Arabic	Bantu	Dravidian	Indo-Aryan	Sino-Tibetan
Listening Comprehension	.36 - .73	.20 -.99			
Reading Comprehension	.20 - .80	.19 – 1.00	.43 - .84	.19 - .85	.25 - .78

To determine the level of difficulty by question type only questions with matching data were examined. Question difficulty ranges from very difficult (.19) to easy (1). Although the level of difficulty of the reading comprehension questions was variable, the majority of the listening comprehension questions fell in the moderate range. Note that the questions are translated so only the question type is common across all languages. The questions are from Arabic, Cinyanja, and Kinyarwanda EGRA assessments. The first three sets of questions are literal and why questions are inferential.

Table 6. Range of difficulty by reading comprehension question type across languages

Question	Difficulty
Who had two chickens?	.91
Who was in Grade 2?	.19
Who did Masi cross the bridge with?	.39
Who did Zaid meet in the garden?	.72
What did Karekezi want to do for them?	.79
After inviting other children from the neighborhood, what did Karekezi give them?	.71
What names did Karekezi choose?	.77
What did Masi like to do at school?	.59
What did Zaid do with the broom?	.80
Where was Masi's school?	.19
Where did Masi go when he crossed the bridge alone?	.19
Where did Zed's toy fall?	.59
Why did Karekezi sing a song for them?	.68
Why did Zaid feel happy?	.60
Why did Zaid feel sad?	.55

Table 7. Range of listening comprehension question type across languages

Question	Difficulty
Who woke up early in the morning?	.62
What kind of animals were living on the way to the king's house?	.61
What happened to the first two lion when the musician played his banjo?	.37
What did the third lion do when the musician played the banjo?	.32
What did the children gather from the field?	.36
In which season did the story take place?	.73
Where did Abu Saeed want to go?	.61
Where was the singer going?	.51
How did the children get up?	.54
Why did the third lion not hear the banjo of the musician?	.20

Note: The following descriptors were used: above .75 is considered easy; between .25 and .75 is considered moderate, and below .25 is considered difficult.

To better illustrate the relationship between a text and the questions, and the range of difficulty, Appendix B includes three passages, and the questions associated with each. The two Bantu passages demonstrate the influence of sample dependence. The first passage is simple, four of the questions are literal, and the answers are explicitly found in the text, yet the questions are in the moderate and difficult range. The second passage is more complex, but four of the questions were also literal with answers explicitly found in the text. All of these questions would be considered easy. This indicates that although text and question type are important, the sample may add greater variability in determining item difficulty.

Table 8. test Blueprint

	Low p-values (most difficult)			Mid-range p-values (medium difficult)			High p-values (least difficult)			Total
	.1-.19	.2-.29	.3-.39	.4-.49	.5-.59	.6-.69	.7-.79	.8-.89	.9-1	
C2.1			3			2	1			6
C3.1										0
D1.1 Sounds	1			1						2
D1.1 Syllables	1		2			2	3	10	5	23
D1.2	1		1		1	5	8	6		22
D2.1					1			1		2
R1.2	3		1		2		4	1	1	12
Total	6	0	7	1	4	9	16	18	6	67

3.2 Benchmarks

A 60% correct on a reading comprehension measure is the minimum needed to report on SDG 4.1.1. To be included, foundational skills must be correlated to a 60% score on reading comprehension. The correlation between foundational skills and reading comprehension is well established. However, because reading is language-based, benchmarks tend to be language specific. Differences in benchmarks across languages is expected due to a number of factors. Examples of language differences are the morphological structure of the language, or the number of symbols. Instructional differences can also account for differences across language groups. For example, when literacy instruction begins, the content of literacy instruction, and the opportunities learners have to practice reading can impact student literacy knowledge.

The table below provides benchmarks for foundational skills associated with a 60% score on reading comprehension.

Table 9. Benchmarks needed to meet 60% on reading comprehension

Reading Component	Arabic	Bantu	Dravidian	Indo-Aryan	Sino-Tibetan
Listening Comprehension	4.0	4.98			
Letter Sounds	52.75	25.76			
Syllable sounds	33.00	36.09	80.72	94.44	91.09
Non-words	14.15	15.59	30.94	41.96	26.78
Familiar words		25.43	34.39	40.42	32.32
Oral reading fluency		19.36	34.43	51.41	45.98
Reading accuracy	66.07				

4. Conclusion

Learning to read is similar across languages therefore, skills assessed to determine student growth toward reading proficiency are similar. The range in item difficulty and benchmarks within language groups indicates that more work is needed in defining item construction. Despite the differences in item difficulty, there was enough data to identify exemplar items for each of the tasks examined.

Appendix A

	Construct	Minimum proficiency
C2.1	Retrieve explicit information in a short grade level continuous text read to or signed for the learner.	When listening to a short grade-2 level continuous text, retrieve explicit information by direct or close-word matching or by simple synonymous word matching when there is limited competing information. It will generally be in response to a “who”, “What”, “when’, or “where” question.
C3.1	Interpret information in a short grade-level continuous text read to or signed to a learner.	When listening to a short grade-2 level continuous text, make simple inferences to connecting pieces of prominent explicit information when there are multiple clues and limited competing information. This will generally be in response to a “why” or “how” question.
D1.1	Identify symbol/sound/fingerspelling and/or symbol/morpheme correspondences	If the grade level curriculum introduces new symbols accurately say or sign common grade 2-level symbol-sound/fingerspelling and/or symbol morpheme correspondences (language and country-specific)
D1.2	Decode isolated words	Accurately say or sign common isolated grade-level words (language and country specific.
D2.1	Say or sign a grade level continuous text at pace and with accuracy	Accurately say or sign a grade 2 continuous text with a few errors (e.g. no more that 10% of the in the text).
R1.2	Retrieve explicit information in a grade level text by direct or close word matching	Retrieve a single piece of explicit information from a grade 2-level text by direct or close word matching when the information required is adjacent to the matched word, and there is no competing information. It will generally be in response to a “who”, “What”, “when’, or “where” question.

Appendix B

Bantu

Passage	Questions	Question Level of Difficulty
Masie was in Grade 2. Masie's school was on the bank of a river. To get to school, Masie had to cross a bridge. Every day, Masie crossed the bridge with his friend Izeki. One day, Izeki didn't go to school. Masie likes to study and play football at school. Tero, he crossed the bridge to school alone.	1. Who was in Grade 2? Masi	.19
	2. Where was Masi's school? On the banks of the river	.19
	3. Who did Masi cross the bridge with? Izek	.39
	4. What did Masi like to do at school? study and play football	.59
	5. Where did Masi go when he crossed the bridge alone? to school	.19

Bantu

Passage	Questions	Question Level of Difficulty
Karekezi had two very beautiful chickens. One day he wanted to give them names. Karekezi invited the children from the neighborhood and gave them pieces of paper. Then they wrote the names on the pieces of paper. Karekezi chose two very beautiful names. He named one chicken Gakokokazi and the other Gasake. Karekezi sang a song, and they danced.	1. Who had two chickens? Karekezi	.91
	2. What did Karekezi want to do for them? Name them.	.79
	3. After inviting other children from the neighborhood, what did Karekezi give them? Papers	.71
	4. What names did Karekezi choose? Gasake and Gakokokazi	.77
	5. Why did Karekezi sing a song for them? /He was happy; he was happy; He had named his chickens.	.68

Arabic

Passage	Questions	Question Level of Difficulty
Zaid threw his toy, and it fell outside the garden wall. He went to get it and found the floor dirty; He felt sad. He returned to his house and brought the broom, then began cleaning the garden floor from the leaves. The cleaner was cleaning the garden, so he greeted Zaid with a smile. Zaid felt happy and returned home happily.	1. Where did Zed's toy fall? Outside the garden fence.	.59
	2. Why did Zaid feel sad? Because the ground is dirty.	.55
	3. What did Zaid do with the broom? Clean the garden ground from the leaves	.80
	4. Who did Zaid meet in the garden? The cleaner	.72
	5. Why did Zaid feel happy? Because he participated in cleaning the street or because he helped the cleaner	.60

TECHNICAL DOCUMENTATION SUPPORTING EXEMPLAR
EARLY GRADES READING ITEMS

Email:
uis.information@unesco.org
uis.director@unesco.org

uis.unesco.org
@UNESCOstat