Measurement and Research Support to Education Strategy Goal 1

Development and Pilot Testing of Additional Subtasks for the Early Grade Reading Assessment: EGRA 2.0

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International.

Measurement and Research Support to Education Strategy Goal 1

Development and Pilot Testing of Additional Subtasks for the Early Grade Reading Assessment: EGRA 2.0

Prepared for

Office of Education
Bureau for Economic Growth, Education, and Environment (E3)
United States Agency for International Development (USAID)
Arthur Muchajer, Contracting Officer
Penelope Bender, Contracting Officer's Representative

Prepared by
RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

RTI International is a registered trademark and a trade name of Research Triangle Institute.

Table of Contents

List of Tables

Abbreviations

CBM	curriculum-based measures
CVC	consonant-vowel-consonant
CVCe	consonant-vowel-consonant followed by silent e
cwpm	correct words per minute
E3	USAID Bureau for Economic Growth, Education, and Environment
EdData II	Education Data for Decision Making project
EGRA	Early Grade Reading Assessment
GES	Ghana Education Service
LORU	**L**anguage (words that were said in another language); **O**ther (words from other category); **R**epetition (of words); **U**nique (total unique words)
RTI	RTI International (registered trademark and trade name of Research Triangle Institute)
USAID	United States Agency for International Development

1. Background

1.1 Justification for Additional Subtasks

The Early Grade Reading Assessment (EGRA) is an open-source assessment composed of individual subtasks that measure some of the foundational skills needed for reading acquisition. Developed initially in 2006 by RTI International with funding from the World Bank and the United States Agency for International Development (USAID), as of mid-2016 the EGRA had been adapted in over 100 languages for use by more than 50 organizations in over 70 countries (see the website of the Global Reading Network for the 2016 *EGRA Toolkit, Second Edition,* https://globalreadingnetwork.net). It is administered individually in an interaction between a trained assessor and an individual, usually a child in primary school.

Not all of the EGRA subtasks are administered in every country or context. Instead, the EGRA instrument is a flexible template from which users can select the subtasks that will help to answer their research questions. The existing EGRA subtasks are described in *Table 1.* The four noted with an asterisk are recommended as useful for a first-time EGRA survey if little is known about the literacy skills of the population of interest. Each subtask has specific purposes and, like all assessments, each also has limitations (Dubeck & Gove, 2015). For example, the existing subtasks are informative, but refinements and additions could improve the EGRA's ability to measure students' reading comprehension, writing, and oral language skills.

Table 1. Existing EGRA Subtasks

Name	Description
Orientation to Print	Measures knowledge of early print concepts such as a word, letters, and directionality. It is untimed and does not have a discontinuation rule.
Letter Name Identification	Measures knowledge of letter names. 100 letters are presented in random order in both upper and lower case. It is timed to 60 seconds and is discontinued if none of letters in the first line (i.e., 10 letters) is read correctly.
Letter Sound Identification*	Measures knowledge of letter-sound correspondences. 100 letters are presented in random order in both upper and lower case. It is timed to 60 seconds and is discontinued if none of the sounds in the first line (i.e., 10 letters) is produced correctly.
Initial Sound Identification	Measures the ability to discriminate beginning sounds. Three words are presented and the aim is to identify the word that begins with a sound that is different from the other two. It is oral and has 10 sets of words. It is discontinued if no points are earned in the first five items.
Segmentation (phoneme or syllables)	Measures the ability to segment a word into individual phonemes or syllables. This subtask is oral and has 10 items. It is discontinued if no points are earned in the first five items.
Syllable Identification	Measures the ability to read individual syllables. 50 syllables are presented. It is timed to 60 seconds and is discontinued if none of the first five syllables is read correctly.

Name	Description
Familiar Word Reading	Measures the ability to identify individual words from grade-level text. 50 words are presented. It is timed to 60 seconds and is discontinued if none of the words in the first line (i.e., five words) is read correctly.
*Nonword Reading**	Measures the ability to decode individual nonwords, following common orthographic structure, from grade-level text. 50 nonwords are presented. It is timed to 60 seconds and is discontinued if none of the words in the first line (i.e., five words) is read correctly.
*Oral Reading Fluency**	Measures the ability to read a grade-level passage of approximately 60 words. It is scored for accuracy and rate. It is timed to 60 seconds and is discontinued if none of words in the first line (i.e., about 10 words) is read correctly.
*Reading Comprehension (with or without lookbacks)**	Measures the ability to answer questions about the grade-level passage. Questions include explicit and inferential examples, and lookbacks (i.e., referencing the passage for the answer) can be used if appropriate.
Cloze	Measures the ability to identify a word among several choices that would complete the sentence using the correct part of speech. It is untimed and does not have a discontinuation rule.
*Listening Comprehension**	Measures receptive language of an orally read passage with both explicit and inferential questions. It is untimed and does not have a discontinuation rule.
Vocabulary	Measures receptive language skills of individual words and phrases related to body parts, common objects, and spatial relationships. It is untimed and does not have a discontinuation rule.
Dictation	Measures the ability to spell and use grammar in a grade-level sentence. Words can be scored for partial representation.
Interview	Gathers information about the child that is related to literacy and language development (e.g., first language; access to print). It is self-reported by the child.

* Recommended as useful for a first-time EGRA if little is known about the literacy skills of the population of interest.

In 2016, RTI International suggested piloting additional subtasks to enhance the instrument. Researchers from RTI collaborated with a team of researchers in Accra, Ghana, to adapt, refine and pilot four subtasks. The work was funded under an EdData II task order, "Measurement and Research Support to Education Strategy Goal 1." This report describes the skills assessed, the four subtasks, the pilot, and the results.

1.2   Proposed and Revised Work Plan

Our original work plan included two main activities. The first proposed activity was to pilot a word-reading subtask and a sentence choice subtask, to better understand reading comprehension; and to pilot a revised method to score sentence dictation, intended to allow more variation in assessing spelling abilities. The second proposed activity in the original work plan was to pilot an alternate way to present EGRA results. After we submitted the original work plan, additional conversations internally and with the USAID Goal 1 Team resulted in revisions. Through these conversations, we jointly

decided in April 2016 to pilot a subtask to measure expressive language skills, we replaced modifications to the sentence dictation subtask with an individual word spelling exercise, and we removed the alternate way to present EGRA results. The subsequent sections of this report describe the activities in the revised work plan.

2. Skills Assessed

2.1 Reading Comprehension: Word Choice and Sentence Choice

The existing EGRA instrument has two subtasks that measure reading comprehension. It measures the ability to read a grade-level passage orally, and the reader is asked five comprehension questions related to the text read. For readers who complete the passage in the allotted time (i.e., 60 seconds), the assessor asks them all five questions. Slower readers are asked between 0 and 4 questions, depending on the amount of text they attempted. Having just a few items to measure reading comprehension is a criticism of the existing EGRA.

Another minor criticism of the EGRA is that students respond to comprehension questions in an *open-ended format,* in which they provide oral answers based on their understanding of the text. Responses are accepted in multiple languages, not just the language of the reading passage. This approach reduces the need for oral language skills in the language of the assessment, but expressive language abilities remain a confounding construct within the subtask. So even though reading a passage of connected text and answering related questions is considered an authentic education activity, we considered techniques that other researchers have relied on to gauge reading comprehension.

As we considered the alternative formats to measure reading comprehension, we had to consider the complexity of the construct. The purpose of reading is to make meaning from the text. Yet measuring comprehension reliably is beyond the scope of most assessments, as it requires attending to multiple constructs (e.g., background knowledge, vocabulary, motivation, attention) that are difficult to measure in a limited amount of assessment time (Snow & RAND Reading Study Group, 2002; Sweet & Snow, 2003). But to develop a valid reading passage that reduces the role of some of these constructs is challenging. For example, a passage about fishing will be more likely understood by a child who regularly fishes because she knows fishing terminology, and she will be motivated to persist with the passage and can connect ideas that are not directly stated because of her knowledge of the topic (Anderson, Wang, & Gaffney, 2006; McNamara & Kintsch, 1996). Developing a valid passage also relies on other structural issues, such as text cohesion, which requires each individual sentence to contribute to the overall passage. Furthermore, the readability of the text is influenced by sentence structure (i.e., syntax), word choice, and meaning. With these complexities in mind, we considered alternative formats to measure reading comprehension.

One common format is the *cloze procedure* (Taylor, 1953) which requires readers to fill in words that either are missing at regular intervals or are omitted following some other criteria. Critics of cloze are concerned that it does not represent typical reading because the reader is forced to scan and rescan text to determine the missing word. Although some researchers identify cloze as a measure of sentence-level skills (Shanahan, Kamil, & Tobin, 1982), the length of a passage has to be adequate to accommodate missing words, and text cohesion is still required between sentences. Furthermore, requiring readers to supply the missing words burdens their expressive-language skills and therefore does not reduce the constructs assessed.

A similar format, *maze,* also confines comprehension to sentence boundaries and is similar to cloze by omitting a word (Parker, Habrouck, & Tindal, 1992). The difference is that it provides several options to select the missing word, which reduces the use of expressive language and is described by Snow and Sweet (2003) as a challenge to really understanding reading comprehension. To understand how to complete a sentence with a missing word, students need to be familiar with this type of classroom assessment, although it does work well in a group-administered format.

A *multiple-choice* format is another option for capturing sentence-level comprehension. The student reads a sentence and compares three or more responses, and selects the best option. This task also reduces the use of expressive-language skills. Similar to maze and cloze, however, it requires some familiarity with this type of assessment as well as exposure to formal schooling. The developers must carefully create three to four incorrect responses (i.e., foils or distractors) that do not make the actual answer obvious or create two valid answers, for example.

We explored two simple reading comprehension subtasks that measure word- and sentence-level comprehension using a protocol in which all items are attempted. The intention is to give EGRA users confidence that students were measured on comprehension independent of their reading rate (i.e., speed) and that expectations for use of expressive language were minimized.

Subtask 1: Word Choice

We developed a Word Choice subtask that measures word level comprehension. This subtask was adapted from Alcock et al.'s (2000) word-reading task, which mixes real words and fake words (nonwords). The subtask's effectiveness and reliability are based on mixing real words from the language of assessment with orthographically legal nonwords, which "are visually and phonologically similar to real words, so that decoding at the phonemic level is necessary to distinguish them from real words" (Alcock et al., 2000, p. 544). This format helps to detect which readers are only decoding words, and which children are both decoding and understanding what they are reading.

This format was considered because it has been shown to consistently and accurately measure whether students after a few years of schooling are both understanding and decoding the presented words (Alcock et al., 2000). We were curious to know whether

---

the Word Choice subtask would have the potential for use with children at the onset of formal reading instruction (i.e., typically grade 1), to measure literacy growth in lower primary grades. To enable collection of evidence of this growth, the word list included orthographic patterns (e.g., in English, consonant-vowel-consonant [CVC]: "bed") that are common in lower primary reading material. To reduce measurement error, the assessment procedures required the student to evaluate each word as either "real" or "fake" (e.g., "tick the real words and put an X next to fake words"). This design yielded four data points for each item:

- Child correctly identified a real word as real.

- Child correctly identified a nonword word as fake.

- Child incorrectly identified a real word as fake.

- Child incorrectly identified a fake word as real.

Subtask 2: Sentence Choice

We also developed the sentence-reading subtask called Sentence Choice. Similar to the word-reading comprehension subtask, this subtask was also adapted from Alcock et al.'s (2000) work. In this subtask, readers were presented with individual sentences that they could discern were either obviously true or obviously false based on their everyday knowledge (e.g., "Goats fly in the sky"). Reading of independent sentences is considered to be a reliable measure of reading comprehension because the sentences are not part of a cohesive text passage (McNamara, Graesser, & Louwerse, 2012). This subtask was developed to be used with students who could read connected text with grade 2 expectations.

2.2   Spelling: Word Dictation

The existing EGRA instrument has one subtask that measures spelling. It measures the ability to spell a sentence that is spoken (dictated) aloud, with the words scored as either correct or incorrect. Using binary scoring limits the variation in the results and increases the presence of floor effects. This lack of variation is a criticism of the existing EGRA. Plus, it is a missed opportunity to understand early literacy skills.

An extensive body of research supports the role of early spelling in later literacy achievement (Arab-Moghaddam & Sénéchal, 2001; Chiappe, Siegel, & Wade-Wooley, 2002; Ehri, 1989; Ehri, 2000; Ellis & Cataldo, 1990; McBride-Chang & Suk-Han Ho 2005; Morris & Perney, 1984; Torgesen & Davis, 1996). For example, Torgesen and Davis (1996) found that successful nonsense-word spelling was a stronger predictor of reading achievement than a basic phonological awareness task that involved identifying beginning sounds. Furthermore, measures of spelling have been supported as robust predictors of later reading achievement for native speakers and nonnative speakers (Henderson, 1990; McBride-Chang & Suk-Han Ho, 2005). The research in this area demonstrates that spelling and reading both use knowledge of the relationships between

letters and sounds. Because some words may be similar in appearance (e.g., in English: mane, man, men), and memory for the multitude of appearances is limited, readers must understand how letters and sounds correspond. Knowing how to spell a word makes it quickly accessible for automatic reading. Such studies support the inclusion of a spelling task in an early literacy-screening instrument (Ehri, 1989; Ehri, 2000; Ellis & Cataldo, 1990; Estes & Richards, 2002; Invernizzi & Hayes, 2004).

A qualitative approach for scoring words provides a range of scores, which can yield a near-normal distribution. Because of its high correlation to other beginning literacy tasks (such as letter knowledge, phonological awareness, and word reading), it provides a lot of useful information (Invernizzi & Hayes, 2004). Furthermore, a spelling subtask that is scored qualitatively—that is, points are awarded for partial correctness of a word— inherently yields more variation in the data than a simple binary method.

Subtask 3: Word Dictation

To capture the developmental nature of spelling skills, we adapted a spelling measure and scoring protocol that has been used in curriculum-based measures (CBM) as described by Wright (1992). For the Word Dictation subtask, a list of words was read individually in a time-limited manner and the students attempted to spell them. Items (i.e., spelling words) were representative of the texts and expectations for that grade level. They were scored for correct letter sequences, or pairs of letters in a word that were written in the proper sequence. As described by Wright (1992) the beginning and ending of the word were scored for "phantom" spaces so the word "talk," a four-letter English word, would have five possible sequences to be scored:

- Sequence 1: space t

- Sequence 2: t a

- Sequence 3: a l

- Sequence 4: l k

- Sequence 5: k space

Besides yielding more variation in scoring, this methodology has other benefits. First, scoring for partial correctness allows for capturing emergent to beginning writing abilities and growth. Second, scoring just for sequencing, instead of representing particular orthographic patterns, has relevance for languages without extensive research on the order in which they are best learned (e.g., in English, the pattern of *ai* is learned before *eigh* to represent the sound of long *a*).

2.3   Oral Language: Semantic Fluency

The existing EGRA contains two subtasks for measuring receptive oral language skills. In the first, the listening comprehension subtask, the student listens to a short story of approximately 30 to 50 words, depending on the orthographic structure of the language,

and answers explicit and inferential question. This subtask is helpful to explain results on the reading comprehension subtasks (Durrell, 1969). Yet this item produces limited information, as it has only a few items.

The second existing EGRA measure for oral language is a vocabulary subtask. It measures receptive language of individual words and phrases by asking pupils orally to identify human body parts, recognize objects in the testing environment, and respond to simple commands.

Both of these receptive-language subtasks measure the skills that are used in listening and reading, and are considered reliable to score. Even so, we were interested in gathering more refined language information.

Oral language skills serve as the foundation for literacy acquisition. Recognizing the language skills a reader brings to the text helps to understand the assets that they can apply to the task of reading unfamiliar text. Some students might be able to read (i.e., decode) individual words correctly, but without the relevant vocabulary and background knowledge, they will not understand the text (Gough & Tunmer, 1986). Conversely, students with strong oral language skills can understand more.

Word knowledge can be measured both *receptively* and *expressively,* and such measures are often compared. Examining meaning at the word level provides useful insight into reading skills (Nagy & Herman, 1987). A common assessment to measure receptive language is for the student to point to a picture that represents one (out of four) orally presented words, as is done in the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007). The picture format can also be used to have the student supply a label or a synonym for a particular picture. Using pictures provides specificity for a concept (e.g., "Show me the picture for the word 'giggle,'"). The main challenge of a picture format is to gather images that are familiar in the assessment context.

Measuring students' ability to describe what they see in a complex illustration or an activity is a valuable way to measure expressive language skills. This format has the advantage of producing variation in the results, as the students are not limited in their responses. However, this format has disadvantages for children who might be shy or reluctant to speak in front of the assessor, which might lead to some inconsistent administrations.

Subtask 4: Semantic Fluency

We ultimately adapted the subtask called Semantic Fluency to measure verbal functioning (Benton, 1968; Lezak, Howieson, Bigler, & Tranel, 2012). This type of task is used to measure verbal ability, lexical knowledge, and lexical retrieval. This type of task is considered useful and therefore valid because it measures both oral language and attention (i.e., executive control). In this subtask, students were given one minute to retrieve as many unique words as they could, first within a provided semantic category and then words from any category of the student's choosing. The student retrieved words, focused on the category, and avoided repetition. Deficits in verbal ability or

attention will influence performance in reading skills, and therefore this type of task is considered an efficient screening instrument.

3.   Pilots

We adapted, refined, and piloted the four subtasks listed in *Table 2,* in English and Akuapem Twi, over a period of several months. First, we adapted existing measures following a format used with previous EGRA subtasks. Next, in June 2016, in collaboration with the research team in Accra, Ghana, we developed additional items and refined the instructions and procedures during several field tests (i.e., mini pilots). In July 2016, at the end of the academic year, we piloted the three English subtasks alongside two of the existing English EGRA subtasks (nonwords and oral reading fluency with reading comprehension). In September 2016, at the onset of the new school year, we piloted the four subtasks in Akuapem Twi with two existing Akuapem Twi EGRA subtasks (nonwords and oral reading fluency with reading comprehension). This section describes the pilot process.

Table 2. Piloted Subtasks

Name	Description
Word Choice	Measures knowledge individual of word reading. It requires the student to decode and determine if the word has meaning. It is untimed. No discontinuation rule. Individual or group format.
Sentence Choice	Measures ability to understand grade-level sentences. It requires the student to decide if the meaning is realistic or silly. It is untimed. No discontinuation rule. Individual or group format.
Word Dictation	Measures ability to spell common words used in lower primary reading material. Words are scored for partial correctness. Twenty seconds per word for consistency. No discontinuation rule. Individual or group format.
Semantic Fluency	Measures the ability to name as many unique words as possible for a particular category. One to three categories recommended. Each category timed to 60 seconds. No discontinuation rule. Individual format.

The Ghanaian research team based in Accra had extensive experience using EGRA and developing education assessments. This experience made them well-placed to contribute to the adaptation and refinement of the four piloted EGRA subtasks. For example, the research team understood the need to limit amount of time that each child could be assessed to still get valid results. Also, having conducted assessor trainings in the past, they knew the importance of having specific and understandable procedures and instructions. Furthermore, they recognized the issues of assessing in a multilingual environment.

With our Ghanaian colleagues, we followed these steps for each of the four subtasks:

1. Began with an adapted version
2. Discussed its purpose and intended constructs
3. Identified problems with the procedures and the instructions
4. Adjusted the procedures, the instructions, and the content
5. Back-translated
6. Trained assessors
7. Field tested and debriefed
8. Repeated Steps 3–7 multiple times for each subtask
9. Pilot tested (English in July 2016; Akuapem Twi in September)

3.1   Word Choice Refinement

The Word Choice subtask did not perform as we expected. During the adaptation phase, we hypothesized that it would be useful to understand students' word-level comprehension. This hypothesis proved incorrect, however, as explained next.

For this subtask, students were presented with 15 individual real and 15 individual nonwords that represented common orthographic patterns for the language (e.g., bed, rope). The students read each word and noted with a mark whether the word was real or fake. However, the students' difficulty in understanding both the directions and the concept of a nonword contributed to it being an unreliable subtask.

In our first field test, we established instructions to tell the students how to mark their paper. We were open to any phrasing that the students would understand, knowing that a phrase used in Ghana might be unreliable elsewhere. The assessors used the term "tick" in the initial field tests to explain to students how they should note that something was correct. In the final field tests, the term was changed to "correct mark." It took two field tests to determine that the phrase "X mark" was a reliable way to explain to students how to note that something was incorrect. Yet even when they understood the markings, the instructions were not sufficient to help them understand the subtask.

In consultation with the research team, we field tested multiple ways to phrase the directions to explain the subtask. The administration always began with the assessor doing an example on the chalkboard while thinking aloud to decide if the example word was real or fake. Thereafter, the students would do several guided examples before starting the subtask. We tried real/fake; real/not real; I know this word/I don't know this word; This word has meaning in English/The word has no meaning in English. We tried administering this subtask individually and in groups of six students. None of these configurations helped the students to understand the subtask, as can be seen in the results (see Results section of this report).

If the instruction challenges can be overcome, the simplicity of compiling the content for Word Choice makes this an appealing measure of word-level comprehension. The real

words we used had common grade-level orthographic patterns[1] and represented objects, actions, persons, and descriptions familiar to students in lower-primary grades. The nonwords had the same orthographic patterns as the real words.

Ultimately, the administration format that we selected for the piloting of the Word Choice subtask was to convene a group of six students, with each student reading the words from their own paper and responding on paper. We created two versions of the pupil response booklet by scrambling the order of the items. Later, the students' responses were entered electronically into the Tangerine® software for immediate analysis.

3.2   Sentence Choice Refinement

The Sentence Choice subtask performed better than we expected. During the adaptation phase, we hypothesized that it would be useful to understand reading comprehension with students in grade 2 and not useful with grade 1 students. Yet our field tests in grades 1 and 2 and the pilot in grade 2 suggested it is valid for both levels.

In this subtask, the students were presented with 10 sentences that were obviously true and 10 sentences that were obviously false (i.e., silly; not true). The markings and phrases to identify whether the sentence was true or false were consistent with those for the Word Choice subtask (i.e., make a correct mark; mark an X). Other similarities in administration included providing practice items, field testing it in a one-to-one format and in a group of six students, and creating two forms to minimize copying from a neighbor.

The sentences had between three and five words. All of them described an object or the use of an item familiar to students in lower primary (e.g., animals, numbers, colors). To expose the prevalence of guessing, each false sentence had a paired true sentence that varied by one word (e.g., Dogs lay eggs. / Chickens lay eggs).

The Sentence Choice subtask that we piloted was administered in a small group of six students, with each student reading the sentences from their own paper and responding on their paper. Later, the students' responses were entered electronically into the Tangerine software for immediate analysis.

3.3   Word Dictation Refinement

The Word Dictation subtask performed as we expected. In Ghana, calling the subtask "spelling" would have been interpreted as the student orally spelling the words, so we used "word dictation" instead.

The refinement in the field test consisted of experimenting with the amount of time between each word and replacing a couple of words so they did not need a sentence to be heard correctly. We did this to save assessment time and to avoid the students writing words that were in the assessor-provided contextual sentence instead of the

---

[1] For English, the orthographic patterns were CVC, r-controlled, CVCe, diphthongs, nasals, and common long-vowel patterns.

target word. For example, we removed the word "sing," which was confused with the word "sink" because they are minimal pairs—the final sound differs by voiced/unvoiced.

The Word Dictation subtask that we piloted was administered in a small group of six students, with the assessor saying each word aloud and each student responding on paper. Later the students' responses were entered electronically into the Tangerine software for immediate analysis.

3.4   Semantic Fluency Refinement

The Semantic Fluency subtask was well received by students and assessors. It was the only subtask of the four that we developed that has to be administered individually.

In our initial adaptation, we considered several criteria for the categories for which students could be asked to supply examples. First, the categories should be familiar to students in lower primary. Second, the categories should not have a lot of interference from another language. This related to a third criterion: that the assessors should be able to score the subtask responses reliably.

Through our field tests, we determined that it would be helpful to classify the categories of items, which would then help to determine the construct assessed:

- A *limitless* category was one that the students knew well in the language of the assessment, such that they could attain at least 10 items in that language, but reaching 20 items in 60 seconds would not be unusual. An example of a limitless category in Akuapem Twi was "body parts," because students begin to learn body parts before they enter school. A limitless category in English was "animals," because in this context, students could identify farm animals, pets, and animals from distant places.

- A *constrained* category was one that was limited by the nature of that category. For children in lower primary, they would most likely be able to identify only 10 or fewer items from that category. Examples of constrained categories were people in a family, colors, or a fruit.

During the assessment, the assessor would introduce one category for the students to respond to. For a second round, the students could choose words from any category they liked.

We noted that the testing environment influenced responses for Semantic Fluency. When we assessed in classrooms that had print on the walls, for some students, these visuals would have priming effect and contribute to the student's next utterance. This occurred both when there was print related to the provided category and when the students were allowed to choose "any category." To reduce these priming effects, we had students sit with their backs to the print.

We field tested several scoring procedures. We determined that responses that were phrases but represented a single idea would be counted as one item. Essentially the

students were penalized by saying too much during the timed portion. Examples that we heard included "white board," "hand sanitizer," and "black dog". Also, responses that varied by one word were scored separately. For example, when students were told to choose "any category," the student who said "white crayon, black crayon, red crayon" was given three points.

We also tried different methods for totaling the points, because we noted it was taking the assessors between four and five minutes per student to score two categories—not including the four minutes that it took to administer them. To try different totaling methods, we created hypothetical lists for all the assessors to score and we developed a scoring method.

At the bottom of the paper, assessors wrote the acronym LORU: L = Language (words that were said in another language); O = Other (words from other category); R = Repetition (of words); U = Unique (total unique words). Also, as assessors were reviewing the list, they were asked to write the letters L and R next to words from those categories, as words said in another language and repetitions would not be included in the Unique total. The use of the acronym helped to achieve reliability among the assessors on the totals, but it still required two to three minutes to total and another minute to enter the results into Tangerine.

3.5 English Pilot

In July 2016 the new English subtasks were piloted with grade 2 pupils at the conclusion of their academic year at a government school in Accra that uses English as a language of instruction. The pilot followed the multiple field tests in which the procedures, instructions, and content were refined. The English pilot was conducted by the same assessors who had been involved with all of the field testing. Two subtasks, Sentence Choice and Word Dictation, were administered to groups of six students, followed by individual administration of Semantic Fluency, Nonwords, and Oral Reading Fluency with Comprehension. The results of the English pilot appear in Section 4.

3.6 Akuapem Twi Pilot

The English subtasks had procedural changes and adjustments each time they were field tested. Therefore, we delayed piloting the Akuapem Twi version until after the English pilot. Because of the school holiday from mid-July through August, the Akuapem Twi pilot was conducted over four days in September 2016 with grade 3 pupils. Piloting with grade 3 students at the beginning of their academic year increased the likelihood of comparability to the grade 2 students who were in the English pilot in early July. The district was selected because Akuapem Twi is predominantly spoken in the locality and it is also a Ghana Education Service (GES) approved local language of instruction for schools in the district.

The Akuapem Twi pilot followed procedures similar to those of the English pilot. As with English, three of the new subtasks—Word Choice, Sentence Choice, and Word Dictation—were administered to groups of six students, followed by individual

---

administration of Semantic Fluency, Nonwords, and Oral Reading Fluency with Comprehension. The results of the Akuapem Twi pilot are in Section 4.

4. Results

4.1 Results for Word Choice, Both Languages

As previously noted, there were significant complications in the design of the Word Choice subtask. After several small pilot administrations, concerns were also raised about the ability of students to understand the instructions of the task. Accordingly, it was important to carefully analyze data from this subtask to determine whether it should be recommended for inclusion in the final suite of subtasks.

The Word Choice subtask was piloted in three separate ways: (1) 96 students were administered this task along with invented words and sentence choice in English; (2) test-retest procedures were administered for 48 students (on Word Choice and Sentence Choice) in English; (3) 155 students were administered this task alongside all other proposed subtasks in Akuapem Twi. All three of these pilots provided evidence that this subtask (in its current form) is not recommended for further use (particularly in Akuapem Twi).

In both the English and Akuapem Twi pilots, the average performance on Word Choice was the first reason for concern. On average, students were able to correctly answer 60% of the Word Choice questions in English and 45% of the questions in Akuapem Twi. Since this task was designed as a true/false measure, the expectation was that students would receive 50% correct by chance alone. While there was an extremely high probability of scoring 45% by purely guessing, there was also a nearly 20% chance of scoring 60% on a 30-item true/false test. This outcome does not provide confidence that the test was accurately measuring student ability, as opposed to pure guessing on a binary response exam.

Since the Word Choice subtask was designed in part as a decoding task (where students were asked to read a word and state whether they had seen it before), we examined the relationship between the Invented Word task and the Word Choice task. In English, the correlation was 0.54—which was below the traditional threshold of 0.70 and lower than expected. In Akuapem Twi, the relationship was an even weaker 0.14. Its relationship to Sentence Choice was 0.55 for English and 0.51 in Akuapem Twi, suggesting that these two tasks were not measuring the same construct. Furthermore, in Akuapem Twi only, we also piloted this task alongside a reading comprehension measure and found the correlation to be 0.15. This all serves as evidence that the subtask was not measuring the skill(s) that we intended them to measure.

Psychometric item-level analyses also pointed to a particular problem. In the Akuapem Twi assessment, 9 of the 10 most difficult items were "nonwords." In the English assessment, 7 of the 10 most difficult items were "nonwords." While intuitively it makes sense that nonwords may be more difficult, the task was simply designed to have

students mark whether they knew a given word. Nonwords, therefore, were just words that were unknown to the students. The fact that these words proved to be the most difficult shows that students were regularly, incorrectly identifying these nonwords as words that they knew. One explanation from the assessors was that students were either confused by the fact that they were provided with "fake" words by an authority figure, or that many of the nonwords seemed like they could have been real words and therefore confused the students.

Final evidence of this subtask's inconsistency came from the English test-retest pilot. With 48 students taking the same Word Choice subtask twice within the same day, we expected an extremely high correlation between scores from the first and second administration (with potentially higher scores in the second because they remembered their answers, fixed some mistakes, and became more familiar and comfortable in the second administration). On average, students changed 8 of their responses from Test 1 to Test 2 (with just over half being correct changes). In all, 90% of students changed at least one response, and half of the students changed more than 7 responses. Ultimately, while students did perform better on the second administration (58% compared with 55%), the difference was not statistically significant. Exactly half of the students improved their scores from one administration to the next, while the other half saw a decrease in performance. Lastly, the correlation between the two tests was only 0.69—which is far below expectations given that testing occurred on the same day.

4.2   Overall Results of the English Pilot

The English pilot data were collected from 150 grade 2 students who were administered the following subtasks: Semantic Fluency (two categories), Sentence Choice, Word Dictation, Invented Words, Oral Reading Passage, and Reading Comprehension.

The Sentence Choice subtask was scored in two ways: (1) Each sentence was scored independently (scale of 0 to 20); (2) sentences were paired and a point was awarded if a student correctly answered both sentences in the pair (scale of 0 to 10). This latter approach was considered more rigorous and also reduced the guessing parameter from 50% to 25%.

Mean scores for all English subtasks are displayed in *Table 3.*

Table 3. Mean Scores for English Pilot

Subtask	Mean score
Semantic Fluency (Animals)	9.8 words per minute
Semantic Fluency (Any Category)	10.5 words per minute
Sentence Choice (Independent)	65.3%
Sentence Choice (Paired)	43.9%
Word Dictation (Letters)	44.8%
Word Dictation (Words)	24.8%

Subtask	Mean score
Invented Words	7.7 words per minute
Oral Reading Fluency	31.1 words per minute
Reading Comprehension	26.8%

There are several important things to note from Table 3. First, while students were able to name nearly 1 more word per minute on the Semantic Fluency task covering "any category," as compared with animals, this difference was not statistically significant. Second, students performed better on the independent Sentence Choice task than they did with paired scoring. This result is to be expected given the reduction in the guessing parameter—and a mean of 43.9% is very unlikely to occur by chance alone. Accordingly, we have focused on the paired scoring approach for the rest of the analyses. Third, students were able to correctly write approximately half of the letters from the dictated words, but only a quarter of the total words were spelled correctly, on average. This finding points to the importance of measuring spelling in terms of more than just correct/incorrect words, as much information is lost with that approach.

Correlations across variables of interest are displayed in *Table 4.* It is clear from this table that the strongest correlations were among Invented Words, Oral Reading Fluency, Reading Comprehension, and Word Dictation (total words). All pairwise correlations across these subtasks were above the 0.70 threshold. Sentence Choice showed correlations between 0.60 and 0.68 with all of those subtasks.

Table 4. Subtask Correlations for English Pilot

	Invented Words	Oral Reading Fluency	Reading Compre-hension	Word Dictation	Sentence Choice	Semantic Fluency
Invented Words	1					
Oral Reading Fluency	0.84	1				
Reading Comprehension	0.71	0.85	1			
Word Dictation	0.80	0.83	0.73	1		
Sentence Choice	0.60	0.68	0.67	0.64	1	
Semantic Fluency	0.33	0.33	0.41	0.35	0.34	1

The two strongest correlations for Sentence Choice were with Oral Reading Fluency and Reading Comprehension, which makes sense given that this task is a simplified comprehension measure with an important component of reading accuracy (if not

fluency). These measures are below the standard cutoff but do provide evidence of a moderate relationship between these variables. The weakest correlations were all in the row for semantic fluency. This subtask was minimally correlated with the others (though we may expect it to be more highly correlated with more basic EGRA subtasks such as listening comprehension).

One of the most standard measures of test reliability is Cronbach's alpha. The test scale provides a singular estimate of the internal consistency of a test. In other words, how well do the different subtasks work together to measure the same construct? Evaluations using Cronbach's alpha have been done many times on EGRA instruments across the globe, but it was important to determine whether these new subtasks could also fit within the EGRA scale.

The Cronbach's alpha for a test scale containing Invented Words, Oral Reading, Reading Comprehension, Word Dictation, and Sentence Choice was 0.92. This signifies very strong test reliability. The scale was estimated with the inclusion of Semantic Fluency as well, but it produced a lower estimate, and the relationship between Semantic Fluency and the rest of the assessment was very weak.

All of these measures are evidence that the new Sentence Choice and Word Dictation tasks are prime candidates for potential inclusion in the EGRA suite of subtasks, but that the Semantic Fluency subtask does not yet provide enough confidence for its inclusion.

4.3   Results for Sentence Choice, English

In addition to the measures above, we administered test-retest reliability procedures for sentence choice. Specifically, we piloted the same version of this subtask (in English) to a sample of 48 students twice in one day. Estimates from the test-retest administrations yielded a less than clear picture of the subtask's reliability. For example, the correlation between paired scores at the two time points was 0.76. This was higher than the conventional 0.70 cutoff for reliability but lower than would be expected for a test given twice in the same day. Additionally, while there was little change in overall scores between Test 1 and Test 2 (52.5% versus 54.6%), students changed approximately 4.5 answers across the two assessments (out of 20 total), on average—just over half of which (57%) were for the better.

Turning toward more positive evidence, item-level and subtask-level measures from the full pilot data provided strong estimates of reliability and appropriateness for the intended grade level. The test scale for paired sentence scores resulted in a Cronbach's alpha of 0.74—a sign of strong reliability. Rasch item analyses also showed that all items functioned as expected and targeted students across a range of abilities that were appropriate for grade 2.

On a 20-item true/false test, the probability of correctly identifying 14 or more answers by guessing alone is approximately 5%. Since 45% of tested pupils answered at least 14 questions correctly, we have evidence that this subtask measured an aspect of reading ability/comprehension beyond the guessing parameter. This was further supported by

the fact that the average Oral Reading Fluency score for students scoring at least a 14 on Sentence Choice was 53.6 correct words per minute (cwpm), as compared with 13 cwpm for those students scoring below 14. Lastly, while only a small number of students were unable to read a single word of connected text (10 students scored 0 on Oral Reading Fluency), they averaged 28% on paired Sentence Choice (which is strikingly similar to the guessing parameter of 25%).

4.4   Results for Word Dictation, English

One of the advantages of the Word Dictation subtask was not in the administration itself but in the scoring. As previously noted, scoring dictated words in a more nuanced fashion than the traditional "correct versus incorrect" for the total word provided valuable information about beginning spellers. For example, the average scores of students across several measures on the Word Dictation subtask are displayed in *Table 5.*

Table 5. Disaggregated Word Dictation Scores (English)

Scoring element	Percent correct
First Letter	70%
Last Letter	53%
Letter Score	45%
Word Score	25%

This table shows that students were able to write correctly approximately 70% of the first letters of words in this task. They were also able to write correctly approximately 53% of the final letters. However, they were able to write correctly only 45% of the total letters and 25% of the total words. This means that it was easier for students to identify the initial letter (or final letter) of a word than it was for them to identify any given letter throughout the subtask.

Even more compelling is the fact that exactly one-fifth of students were unable to spell a single word correctly, but they were still able to correctly write some of the letters. These results are displayed in *Table 6.*

Table 6. Word Dictation Letter Scores by Total Word Scores (English)

Scoring element	Zero words spelled correctly	At least one word spelled correctly
First Letter Score	34%	79%
Last Letter Score	17%	62%
All Letter Score	13%	52%

The data columns in Table 6 display average scores for students who had zero total words correct and those with at least one correct word, respectively. While it is not surprising that students with at least one word spelled correctly would also correctly write more letters, the interesting information comes from the first column. This shows that among students who could not write a single complete word, they were still able to average more than one-third correct first letters in words. They did only half as well with correctly writing the last letters of the words (and slightly worse again as a proportion of total letters). This shows that students begin with the recognition of the first sound/letter in the word (and often recognize the final sound/letter) even when they are unable to spell the entire word correctly.

4.5   Overall Results of the Akuapem Twi Pilot

Data from the 155 students in the Akuapem Twi pilot provided strong evidence that the new subtasks should be revised or reconsidered for use in this language. The average scores displayed in *Table 7* point to the first piece of such evidence, with both Sentence Choice measures (independent and paired) showing means that aligned almost perfectly with scores based on guessing alone. Furthermore, fewer than 10% of students scored 14 or more true/false questions correctly on the Sentence Choice subtask, which is not far from the approximately 6% that would be expected by guessing alone.

Table 7. Mean Scores for Akuapem Twi Pilot

Subtask	Mean score
Semantic Fluency (Body Parts)	9.8 words per minute
Semantic Fluency (Any Category)	6.6 words per minute
Sentence Choice (Independent)	49.2%
Sentence Choice (Paired)	27.1%
Word Dictation (Letters)	34.0%
Word Dictation (Words)	22.8%
Invented Words	3.1 words per minute
Oral Reading Fluency	5.8 words per minute
Reading Comprehension	7.1%

The correlations in *Table 8* further complicate the story. While correlations among the Word Dictation subtask and the three standard EGRA subtasks were moderate, they were significantly lower than in the English pilot. Additionally, the correlations between Sentence Choice and each other subtask were all negative. Although the correlations were small, the negative relationship is confusing and disconcerting.

Table 8. Correlations for the Akuapem Twi Pilot

	Invented Words	Oral Reading Fluency	Reading Compre-hension	Word Dictation	Sentence Choice	Semantic Fluency
Invented Words	1					
Oral Reading Fluency	0.76	1				
Reading Compre-hension	0.76	0.73	1			
Word Dictation	0.68	0.55	0.69	1		
Sentence Choice	-0.18	-0.22	-0.17	-0.18	1	
Semantic Fluency	0.23	0.21	0.30	0.29	-0.11	1

The overall test scale for the assessment (including Invented Words, Oral Reading Fluency, Reading Comprehension, Word Dictation, and Sentence Choice) was 0.79, but the alpha increased to 0.85 when we removed Sentence Choice. This is because the subtask was so minimally correlated with the rest of the test scale. The paired sentences measure on its own yielded a Cronbach's alpha of only 0.51, which is very low for the internal consistency of a single subtask. The Sentence Choice subtask's poor performance likely occurred largely because many students were guessing at random. The question remains whether this happened because the subtask was not appropriate for Akuapem Twi, the directions were not clear, or the piloted students did not have the necessary skills for this task.

There is, however, some promising evidence for the Word Dictation subtask in Akuapem Twi. First, the Cronbach's alpha for the Word Dictation subtask test scale was a robust 0.91. Additionally, *Table 9* shows that even among students who could not write a single complete word correctly, they were still able to write nearly one-third of the first letters of words and one-eighth of the last letters. This means that the Word Dictation subtask had high internal consistency and that the scoring approach provided useful information beyond just word-level spelling.

Table 9. Word Dictation Letter Scores by Total Word Scores (Akuapem Twi)

Scoring element	Zero words spelled correctly	At least one word spelled correctly
First Letter Score	31%	64%
Last Letter Score	12%	47%
All Letter Score	13%	45%

5.   Recommendations

Our recommendations for the next steps with these four subtasks are based on several criteria. First, the subtask should give useful information. To determine whether this was the case, we examined the data from the pilots conducted in July and September. Second, the subtask should be developmentally appropriate for students in lower primary. To determine that benchmark, we considered our experiences with the field testing and the students' responses to the activities. Third, the subtask should be adaptable to other languages with a level of effort that aligns with the available resources in the context in which EGRA is used.

The research team's recommendations by subtask appear below.

5.1   Word Choice

We do not recommend Word Choice for use in lower primary grades. As described in the previous sections, despite our efforts to make it more comprehensible, students could not understand the subtask. The students' inability to grasp what was wanted of them may have contributed to the subtask's seeming inability to be measure the intended construct. Yet as Alcock and colleagues (2000) showed, this subtask has potential for upper grades. *Table 10* outlines our recommendations for this subtask. The pilot versions are in the Annex.

Table 10. Word Choice Recommendations

Word Choice	Recommendations
Use	• Do not use with students in lower primary (grades 1–3). • Further explore potential for grades 4 and higher.
Item Adaptation	• Choose words that are representative of common orthographic patterns for the grade or reading level of interest. • Use 30 words, 12 real and 18 fake (nonwords). • Involve a language expert familiar with early grade reading.
Administration	• Clarify use of markings and terminology to indicate whether the word is real or fake. • Begin with example items on the chalkboard and in the student response book. • Use a small enough group that the assessor has proximity to each student to provide guidance and touch each student's paper. • Limit the assessment time to 3 minutes. • Create two forms, with the same words but in a different order, to decrease effects of copying.
Scoring	• Instruct assessors to enter the student's response exactly, not record whether the response was correct or incorrect. This will help to avoid interference (i.e., similar to a Stroop effect).

5.2   Sentence Choice

At the onset of this activity, one of our stated goals was to "produce one or two simple measures of reading comprehension to complement or replace the existing comprehension measure." Sentence Choice met that goal. This subtask was understood by children in lower primary, the English version had a strong relationship (0.67) to existing reading comprehension measures, and the scores were stable across two time points (0.76). However, the Akuapem Twi version of Sentence Choice did not perform well and we recommend further exploration on the sentence structure. *Table 11* outlines our recommendations for this subtask. The pilot versions are in the Annex.

Table 11. Sentence Choice Recommendations

Sentence Choice	Recommendations
Use	• Use with students in lower primary (grades 1–3).
Item Adaptation	• Create sentences of similar length and structure that are representative of the grade or reading level of interest. • Use 10 pairs of sentences. Each true sentence should have a false pair (Birds fly in the sky. / Goats fly in the sky.) • Involve a language expert familiar with early grade reading.
Administration	• Clarify use of markings and terminology to indicate whether the sentence is true or false. • Begin with example items on the chalkboard and in the student response book. • Use a small enough group that the assessor has proximity to each student to provide guidance and touch each student's paper. • Limit the assessment time to 5 minutes. • Create two forms with sentences scrambled to decrease effects of copying.
Scoring	• Instruct assessors to enter the student's responses exactly, not record whether the response was correct or incorrect. This will help to avoid interference (i.e., similar to a Stroop effect).

5.3   Word Dictation

Another goal of this activity was to find a simple way to get a distribution of spelling abilities that corresponds to beginning reading. The Word Dictation task that we adapted from Wright's (1992) curriculum-based measure would help to fill that need. The subtask was reliable, the results had variability, and it was able to capture students' earliest representations of the written word. *Table 12* outlines our recommendations for this subtask. The pilot versions are in the Annex.

Table 12. Word Dictation Recommendations

Word Dictation	Recommendations
Use	• Use with students in lower primary (grades 1–3).
Item Adaptation	• Select words that are representative of common orthographic patterns for the grade or reading level of interest. • Ensure that when said in isolation, the spelling words will not be confused with other words. For example, avoid minimal pairs that differ by just one sound (e.g., sing/sink). Do not use homophones (ate/eight). • Use 12 words. • Involve a language expert familiar with early grade reading.
Administration	• Begin with example items on the chalkboard and in the student response book. • Use a small enough group that the assessor has proximity to each student to provide guidance and touch each student's paper. • Limit the assessment time to 20 seconds per word. • Provide a pupil response sheet with numbered lines.
Scoring	• Provide a space to score each sequence as well as including a "Word Correct" scoring option. For example, for the word "bed": _b be ed d_ Word Correct

5.4   Semantic Fluency

This subtask was positively received by students and assessors. However, it was minimally correlated with the other EGRA subtasks. We did not examine its relationship to the EGRA listening comprehension or vocabulary subtasks, which would be helpful. *Table 13* outlines our recommendations for this subtask. The pilot versions are in the Annex.

Table 13. Semantic Fluency Recommendations

Semantic Fluency	Recommendations
Use	• Use with students in lower primary (grades 1–3). • Pilot alongside of listening comprehension or vocabulary subtasks.
Item Adaptation	• Develop constrained and limitless categories for the language of interest. • If assessing in multiple languages, use constrained or limited categories in both languages. Use one of these. • Use "any category" as a comparison.
Administration	• Begin with example items on the chalkboard. • Administer one-to-one (assessor with student). • Allow a maximum of 60 seconds per category. • Position it as the last task of the EGRA administration. • Write responses as student utters them.
Scoring	• Use the LORU labeling method on paper before entering the totals onto paper or into Tangerine (see Section 3.4).

References

Alcock, K. J., Nokes, K., Ngowi, F., Musabi, C., Mbise, A., Mandali, R., Bundy, & Baddeley, A. (2000). The development of reading tests for use in a regularly spelled language. *Applied Psycholinguistics*, *21*(4), 525–555. http://dx.doi.org/10.1017/S0142716400004069

Anderson, R. C., Wang, Q., & Gaffney, J. S. (2006). Comprehension research over the past three decades. In K. A. Dougherty Stahl & M. C. McKenna (Eds.), *Reading research at work: Foundations of effective practice* (pp. 275–283). New York: The Guilford Press.

Arab-Moghaddam, N., & Sénéchal, M. (2001). Orthographic and phonological processing skills in reading and spelling in Persian/English bilinguals. *International Journal of Behavioral Development*, *25*(2), 140–147. http://dx.doi.org/10.1080/01650250042000320

Benton A. L. (1968). Differential behavioural effects in frontal lobe disease. *Neuropsychologia 6,* 53–60. https://dx.doi.org/10.1016/0028-3932(68)90038-9

Chiappe, P., Siegel, L. S., & Wade-Woolley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading*, *6*(4), 369–400. http://dx.doi.org/10.1207/S1532799XSSR0604_04

Dubeck, M. M., & Gove, A. K. (2015). The Early Grade Reading Assessment (EGRA): Its theoretical foundation, purpose, and limitations. *International Journal of Educational Development*, *40*, 315–322. http://dx.doi.org/10.1016/j.ijedudev.2014.11.004

Dunn, D. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test: Manual* (4th ed.) Pearson Education.

Durrell, D. D. (1969). Listening comprehension versus reading comprehension. *Journal of Reading*, *12*(6), 455–460.

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities*, *22*(6), 356–365. http://dx.doi.org/10.1177/002221948902200606

Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders*, *20*(3), 19–36. http://dx.doi.org/10.1097/00011363-200020030-00005

Ellis, N., & Cataldo, S. (1990). The role of spelling in learning to read. *Language and Education*, *4*(1), 1–28. http://dx.doi.org/10.1080/09500789009541270

Estes, T., & Richards, H. (2002). Knowledge of orthographic features in Spanish among bilingual children. *Bilingual Research Journal*, *26*(2), 295–307. http://dx.doi.org/10.1080/15235882.2002.10668712

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. http://dx.doi.org/10.1177/074193258600700104

Henderson, E. (1990). *Teaching spelling* (2nd ed.). Boston: Houghton Mifflin.

Invernizzi, M., & Hayes, L. (2004). Developmental spelling research: A systematic imperative. *Reading Research Quarterly*, *39*(2), 216–228. http://dx.doi.org/10.1598/RRQ.39.2.4

Lezak, M., Howieson, D., Bigler, E., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.) New York: Oxford University Press.

McBride-Chang, C., & Suk-Han Ho, C. (2005). Predictors of beginning reading in Chinese and English: A 2-year longitudinal study of Chinese kindergartners. *Scientific Studies of Reading*, *9*(2), 117–144. http://dx.doi.org/10.1207/s1532799xssr0902_2

McNamara, D. S., Graesser, A. C., & Louwerse, M. M. (2012). Sources of text difficulty: Across genres and grades. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 89–116). Lanham, MD: Rowman & Littlefield Education.

McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*(3), 247–288. http://dx.doi.org/10.1080/01638539609544975

Morris, D., & Perney, J. (1984). Developmental spelling as a predictor of first-grade reading achievement. *The Elementary School Journal*, *84*(4), 441–457. http://dx.doi.org/10.1086/461375

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19–36). Hillsdale, NJ: Erlbaum.

Parker, R., Hasbrouck, J., & Tindal, G. (1992). The maze as a classroom-based reading measure: Construction methods, reliability, and validity. *The Journal of Special Education*, *26*(2), 195–218. http://dx.doi.org/10.1177/002246699202600205

Shanahan, T., Kamil, M. L., & Tobin, A. W. (1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly,* 17(2), 229. https://doi.org/10.2307/747485

Snow, C., & RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension.* Research prepared for the Office of Educational Research and Improvement (OERI), U.S. Department of Education. Santa Monica, CA: RAND Corporation. http://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1465.pdf

Sweet, A. P., & Snow, C. E. (2003). *Rethinking reading comprehension: Solving problems in the teaching of literacy*. New York: Guilford.

Taylor, W. (1953). Cloze procedure: A new tool for measuring readability. *The Journalism Quarterly*, *30*, 414–433.

Torgesen, J. K., & Davis, C. (1996). Individual difference variables that predict response to training in phonological awareness. *Journal of Experimental Child Psychology*, *63*(1), 1–21. http://dx.doi.org/10.1006/jecp.1996.0040

Wright, J. (1992). "Curriculum-based measurement: A manual for teachers." *Syracuse City Schools.* Retrieved from http://www.jimwrightonline.com/pdfdocs/cbaManual.pdf

Annex: Piloted Subtasks – 2016 Instruments

English Assessor Protocol

Early Grade Reading Assessment 2.0 Pilot Subtasks
Administrator Instructions and Protocol
English

General instructions

Establish a playful and relaxed rapport with the child through a short conversation (See example topics below). The child should perceive the assessment almost as a game to be enjoyed rather than a test. Use this time to identify in what language the child is most comfortable communicating. Read aloud slowly and clearly ONLY the sections in boxes.

> **Hello. My name is ____ and I live in _____. I'd like to tell you a little bit about myself.** [Number and ages of children; favourite sport, radio or television program, etc.] **1. What do you like to do when you are not in school?** [Wait for response; if student is reluctant, ask question 2, but if they seem comfortable continue to verbal consent]. **2. What games do you like to play?**

Verbal Consent: Read the text in the box clearly to the child.

> - **Let me tell you why I am here today. I work with the Ministry of Education and we are trying to understand how children learn to read. You were picked by chance.**
> - **We would like your help in this. But you do not have to take part if you do not want to.**
> - **I will ask you to read words and sentences, write some words and give ideas.**
> - **Using this stopwatch/device, I will see how long it takes you to do some of the activities.**
> - **This is NOT a test and it will not affect your grade at school.**
> - **I will NOT write down your name so no one will know these are your answers.**
> - **Once again, you do not have to participate if you do not wish to. Once we begin, if you would rather not answer a question, that's all right.**
> - **Do you have any questions? Are you ready to get started?**

Check box if verbal consent is obtained: ☐ **YES**

(If verbal consent is not obtained, thank the child and continue to the next child, using this same form)

A. Date of Assessment:	|___|___| |___|___| |___|___| Day month year
B. Assessor's Name:	
C: Assessor's Code:	|___|___|
C: District:	
D. School Name:	

F. Purpose:	○ 1= Pilot ○ 2 = Technical Adequacy
G. Class:	○ 1 = P1 ○ 3 = P3 ○ 2 = P2 ○ 4 = P4
H: Class Name: (stream)	
I. Pupil Age:	|____|____|
J. Pupil's Gender:	○ 1 = boy ○ 2 = girl
K. Time:	Start |__|___|:|__|___| Finish|__|___|: |__|___|

Subtask: Word Choice	🔖 Children write in booklet.	⏱ 3 minutes

🔈 **Please put your pencils down and listen to me.** [Write a large ✔ and an X on the board. Ask the children what they call these symbols. Use those words throughout the subtask.] **This paper has a list of English words.** [Point to the page with words.] **Some of these words have meaning in English and they are words you know. Some of these words do not have meaning in English and they are words you do not know. You will read them to yourself and decide. If you know the word put a correct mark (✔) in the small box next to the word. If you do not know the word, put a wrong mark (x) in the small box next to the word. Let's do a few examples together.**

[Write the word 'zam' with a box beside it on the board.] **For example, this first word is 'zam'** [Point to 'zam'.] **I don't know this word in English so I put a wrong mark (x) in the box.** [Put a wrong mark (x) in the box.] **Put your finger on the word 'zam' at the top of your paper.** [Check to see that they are pointing to 'zam'.] **Now put a wrong mark (x) in the box next to 'zam'.** [Check to see if they put a wrong mark (x) next to 'zam'.]

[Write the word 'sit' with a box beside it on the board.] **Let's do another example. Put your finger on the word 'sit' on your paper and decide if you know the word. Mark your answer.** [Check that they all marked an answer.]

It is now time for you to do this on your own. When I say 'Start work' begin from number 1 and continue to number 30. Put your finger on box number 1. [Check to see if they are in the correct spot.] **Read each word and decide if you know the word or not. What do we write if we do not know the word?** [Give children a chance to respond. Correct as needed.] **What do we write if we do know the word?** [Give children a chance to respond. Correct as needed.] **You have 3 minutes. Let's get ready. Start work.** [Start timer. Monitor children. At 3 minutes, stop the task.]

Word		Child response*		Word		Child response*	
1.	bed	I know	I don't know	16.	kick	I know	I don't know
2.	horn	I know	I don't know	17.	soud	I know	I don't know
3.	noil	I know	I don't know	18.	thin	I know	I don't know
4.	jump	I know	I don't know	19.	third	I know	I don't know
5.	ploat	I know	I don't know	20.	rill	I know	I don't know
6.	mamp	I know	I don't know	21.	bright	I know	I don't know
7.	mirl	I know	I don't know	22.	wait	I know	I don't know
8.	rope	I know	I don't know	23.	dream	I know	I don't know
9.	sar	I know	I don't know	24.	yan	I know	I don't know
10.	coach	I know	I don't know	25.	tream	I know	I don't know
11.	bim	I know	I don't know	26.	bape	I know	I don't know
12.	shors	I know	I don't know	27.	hill	I know	I don't know
13.	froil	I know	I don't know	28.	shout	I know	I don't know
14.	vight	I know	I don't know	29.	gick	I know	I don't know
15.	baim	I know	I don't know	30.	tham	I know	I don't know

* On this protocol, do not mark if the child was correct or incorrect. Just note the child's response.

Good effort! Let's go to the next page.

Subtask: Sentence Choice	🔖 Children write in booklet.	⏰ 5 minutes

🗣 **Please put your pencils down and listen to me.** [Write a large ✔ and an X on the board. Ask the children what they call these symbols.] **This paper has a list of English sentences.** [Point to the page with sentences.] **Some sentences are <u>TRUE</u> and some are <u>FALSE</u>. You will read them to yourself and decide. If the sentence is <u>true</u>, put a correct mark (✔) in the small box next to the sentence. If the sentence is <u>false</u>, put a wrong mark (x) in the small box next to the sentence. Let's do a few together.**

[Write the sentence 'Goats live in the sky' with a box beside it on the board.] **For example: 'Goats live in the sky'** [Point to the sentence.] **Do goats live in the sky? No. This is false, because goats do not live in the sky, so we will put a wrong mark in the box.** [Write a wrong mark (x) in the box.] **Put your finger on the sentence 'Goats live in the sky' at the top of your paper.** [Check to see that they are in the correct spot.] **Now put a wrong mark in the box next to 'Goats live in the sky'.** [Check to see if they put a wrong mark (x) next to the sentence. Correct as needed.]

[Write the sentence 'Cats have tails' with a box beside it on the board.] **Let's do another example. 'Cats have tails'** [Point to the sentence.] **Do cats have tails? Yes. This sentence is true, because cats have tails, so we put a correct mark (✔) in the box.** [Write a correct mark (✔) in the box.] **Put your finger on the sentence 'Cats have tails' at the top of your paper.** [Check to see that they are in the correct spot.] **Now put a correct mark (✔) in the box next to 'Cats have tails'.** [Check to see if they put a correct mark (✔) next to the sentence. Correct as needed.]

It is now time for you to do this on your own. When I say 'Start work' begin from number 1 and continue to number 20. Put your finger on box number 1. [Check to see if they are in the correct spot.] **Read each sentence and decide if it is true or false. What do we write if the sentence is true?** [Give children a chance to respond. Correct as needed.] **What do we write if the sentence is false?** [Give children a chance to respond. Correct as needed.] **When I tell you to stop, put your pencil down. You have 5 minutes. Let's get ready. Start work.** [Start timer. Monitor children. At 5 minutes, stop the task.]

Sentence		Child response*		Sentence		Child response*	
1.	Dogs lay eggs.	True	False	11.	Bees make honey.	True	False
2.	Chickens have pencils.	True	False	12.	A mouse has a door.	True	False
3.	Six is a shape.	True	False	13.	We eat food.	True	False
4.	A house has a door.	True	False	14.	Red is a colour.	True	False
5.	Football is a game.	True	False	15.	Chickens have feathers.	True	False
6.	Men can walk.	True	False	16.	Boats have tyres.	True	False
7.	We eat paper.	True	False	17.	Six is a number.	True	False
8.	Bees make bread.	True	False	18.	Desks can walk.	True	False
9.	Cars have tyres.	True	False	19.	Birds lay eggs.	True	False
10.	Ball is a colour.	True	False	20.	Football is a meal.	True	False

*On this protocol, do not mark if the child was correct or incorrect. Just note the child's response.

Good effort! Let's go to the next page.

| Subtask: Word Dictation | | | 🔖 Children need ruled paper | 🕐 4 minutes |

🔊 **Put your pencils down and listen to me. Here is a page with lines numbered 1 – 12.** [Point to the paper.] **We will write some English words on these lines. I'll go first. The word I want to write is 'big'.** [Write the word 'big' on the board as you say it.] **b - i - g, big.**

Now you will write some words on your own. If you don't know all of the letters in that word that is okay. You can write just the letters you hear in the words.

Let's practice. Put your finger on the first line. [Check to see that all the children are on the line labeled Example.] **Write the word, 'sat'.** [Wait for the children to write the word 'sat' on the first line. If a child does not start, say, **Just write the letters that you know in the word. It is okay if you don't know them all.**

After the children have attempted the practice word say, **I can tell you are trying. Let's do more. When I say 'Begin' start from number 1 and write the word you hear.** [Check to see that all the children are on number 1.] **If you have not finished a word before I say the next word, that is okay, just start writing the next word on the next line. You will have 20 seconds to write each word. When I tell you to stop, put your pencil down.**
[Read each word. Repeat at 3 seconds. At 10 seconds say the word a third time. Stop the activity after 4 minutes.]

No.	Time	Word						
1	:00	mat	_m	ma		at	t_	Word correct
2	:20	job	_j	jo		ob	b_	Word correct
3	:40	bell	_b	be	el	ll	l_	Word correct
4	:60	sick	_s	si	ic	ck	k_	Word correct
5	1:20	when	_w	wh	he	en	n_	Word correct
6	1:40	late	_l	la	at	te	e_	Word correct
7	2:00	soap	_s	so	oa	ap	p_	Word correct
8	2:20	out	_o	ou	ut	t_		Word correct
9	2:40	sing	_s	si	in	ng	g_	Word correct
10	3:00	near	_n	ne	ea	ar	r_	Word correct
11	3:20	stick	_s	st	ti	ic	ck	k_	Word correct
12	3:40	fight	_f	fi	ig	gh	ht	t_	Word correct

* Cross out the boxes that do not match the way the child wrote the word. If the entire word is correct, circle word correct.

Subtask: Semantic Fluency	🖹Assessor writes in blank page in pupil booklet	⏱ 60 seconds

🕊 **For last activity, we will just talk. What do you like to talk about with your friends and family?** [Wait for response.]

Subtask 1a: **I like animals. We are going to talk about animals. I want you to say as many animals in English as you can until I say stop. You can say any type of animal in English and I will write what you say. Do you understand?** [If the student says no, repeat instruction.] **Start.**

Start the timer. Write down all of the animals that the child says. After 60 seconds say, **Stop. I can tell you are trying.**

Subtask 1b. **I also like to talk about parts of the body. For example: This is my hand.** [Show your hand.] **I want you to name as many parts of the body in English as you can until I say stop. You can say any part of the body in English and I will write what you say. Start.**

Start the timer. Write down all the parts of the body the child says. After 60 seconds say, **Stop. I can tell you are trying. Let's do one more talking activity.**

Subtask 1c. **We just talked about animals and parts of the body. Now I want you to name as many single English words about ANYTHING that you know until I say stop. You can say any single English word about anything. Do you understand?** [If the student says no, repeat instruction.] **Start.**

Start the timer. Write down all of the words that the child says. After 60 seconds say, **Stop. I can tell that you are trying.**

Scoring note to assessor: After the child completes the task, ensure that you can read what you have written. For each topic, use the acronym LORU. 1) Write the letter L next to words said in another language. 2) Write the letter O next to words that are off topic. 3) Write the letter R next to words that are repeated. 4) Total the L words and then cross them off. 5) Total the O words and then cross them off. 6) Total the R words and then cross them off. 7) The remaining words will be recorded as the total unique words (U) for that category.

Subtask 1a (Animals)	Subtask 1b (Body parts)	Subtask 1c (Any category)
1. Total words in another language |___|___|___|	1.Total words in another language |___|___|___|	1. Total words in another language |___|___|___|
2. Total words off topic |___|___|___|	2. Total words off topic |___|___|___|	2. Total words off topic Off Topic is not relevant and not scored for Any Category
3. Total repetitions |___|___|___|	3. Total repetitions |___|___|___|	3. Total repetitions |___|___|___|
4. Total unique words |___|___|___|	4. Total unique words |___|___|___|	4. Total unique words |___|___|___|

Side column:

Start the timer after explaining the task.

✋ When the timer reaches 0, say "stop".

If the child gives you a response in a language other than English, say **Please tell me in English**. This prompt may be given once per category during the subtask.

Good effort! We are done, congratulations!

English Pupil Response Booklet

Early Grade Reading Assessment 2.0 Pilot Subtasks
Pupil Answer Booklet

One pupil answer booklet should be given to each pupil to write his/her responses.

ENGLISH

FORM A

Pupil ID: _____

Subtask: Word Choice

√ = I know this word   x = I do not know this word

1.	bed		16.	kick	
2.	horn		17.	soud	
3.	noil		18.	thin	
4.	jump		19.	third	
5.	ploat		20.	rill	
6.	mamp		21.	bright	
7.	mirl		22.	wait	
8.	rope		23.	dream	
9.	sar		24.	yan	
10.	coach		25.	tream	
11.	bim		26.	bape	
12.	shors		27.	hill	
13.	froil		28.	shout	
14.	vight		29.	gick	
15.	baim		30.	tham	

Subtask: Silly Sentences

EXAMPLE	√ = TRUE	x = FALSE
Goats live in the sky. ☐		Cats have tails. ☐

√ = true x = false	
1. Dogs lay eggs.	
2. Chickens have pencils.	
3. Six is a shape.	
4. A house has a door.	
5. Football is a game.	
6. Men can walk.	
7. We eat paper.	
8. Bees make bread.	
9. Cars have tyres.	
10. Ball is a colour.	
11. Bees make honey.	
12. A mouse has a door.	
13. We eat food.	
14. Red is a colour.	
15. Chickens have feathers.	
16. Boats have tyres.	
17. Six is a number.	
18. Desks can walk.	
19. Birds lay eggs.	
20. Football is a meal.	

Subtask: Word Dictation

EXAMPLE

1. _____

2. _____

3. _____

4. _____

5. _____

6. _____

7. _____

8. _____

9. _____

10. _____

11. _____

12. _____

Subtask: Semantic Fluency

Akuapem Twi Assessor Protocol

Ghana Early Grade Reading Assessment 2.0 Pilot Subtasks
Administrator Instructions and Protocol
Akuapem Twi

General instructions

Establish a playful and relaxed rapport with the child through a short conversation (see example topics below). The child should perceive the assessment almost as a game to be enjoyed rather than a test. Use this time to identify in what language the child is most comfortable communicating. Read aloud slowly and clearly ONLY the sections in boxes.

Maakye, Me din de_____na mete_____. Mepɛ sɛ mɛka me ho asɛm kakra akyerɛ wo. Hello. My name is ____ and I live in _____. I'd like to tell you a little bit about myself. [Number and ages of children; favourite sport, radio or television program, etc.]

1. Woankɔ sukuu a, dɛn na woyɛ? What do you like to do when you are not in school? [Wait for response; if student is reluctant, ask question 2, but if they seem comfortable continue to verbal consent].

2. Agoru ahorow bɛn na wopɛ? What games do you like to play?

Verbal Consent: Read the text in the box clearly to the child.

- **Ma menka nea enti a mewɔ ha nkyerɛ wo. Me ne Nhomasua Adwuma (Ghana Education Service) na ɛyɛ adwuma. Na yɛpɛ sɛ yehu ɔkwan ahorow a mmofra fa so sua akenkan. Wɔanhwɛ biribi pɔtee bi ho na wɔde paw wo.** Let me tell you why I am here today. I work with the Ministry of Education and we are trying to understand how children learn to read. You were picked by chance.

- **Yebehia wo mmoa wɔ saa dwumadi yi mu. Wopɛ nso a, wubetumi ayi wo ho afi mu.** We would like your help in this. But you do not have to take part if you do not want to.

- **Mɛma wo akenkan nsɛmfua ne ɔkasamu bi, na woakyerɛw nsɛmfua bi, na woakyerɛ me nneɛma bi.** I will ask you to read words and sentences, write some words and give ideas.

- **Mede saa afiri yi (wɔɔkye) bɛhwɛ wo mmere tenten a wode bedi dwuma no awie.** Using this stopwatch/device, I will see how long it takes you to do some of the activities.

- **Eyi nyɛ sɔhwɛ a wɔbɛhwɛ ho akyerɛ wo mmɔdemmɔ wɔ sukuu mu.** This is NOT a test and it will not affect your grade at school.

- **Merenkyerɛw wo din, enti obiara renhu wo mmuae a wode mae.** I will NOT write down your name so no one will know these are your answers.

- **Bio, wompɛ a, wubetumi atwe wo ho afi dwumadi yi mu. Yenya fi ase na wompɛ sɛ wobɛma asɛmmisa bi ho mmuae a, wubetumi agyae.** Once again, you do not have to participate if you do not wish to. Once we begin, if you would rather not answer a question, that's all right.

- **Wowɔ nsɛmmisa bi?** Do you have any questions? **Metumi afi ase? Yebefi ase.** Are you ready to get started?

Check box if verbal consent is obtained: ☐ *YES*

(If verbal consent is not obtained, thank the child and continue to the next child, using this same form)

A. Date of Assessment:	|____|____| |____|____| |____|____| Day month year	F. Purpose	○ 1= Pilot ○ 2 = Technical Adequacy
B. Assessor's Name:		G. Class:	○ 1 = P1 ○ 3 = P3 ○ 2 = P2 ○ 4 = P4
C: Assessor's Code:	|___|___|	H: Class Name: (stream)	
C: District		I. Pupil Age:	|____|____|
D. School Name:		J. Pupil's Gender:	○ 1 = boy ○ 2 = girl
		K. Time:	Start |___|___|:|___|___| Finish|___|___|: |___|___|

✊ **Fa wo pɛnsere no to hɔ na tie me.** [Write a large √ and an X on the board. Ask the children what they call these symbols again. Use those words throughout the subtask.] **Ɔfa ha yɛwɔ Twi nsɛmfua bi wɔ ha.** [Point to the page with words.] **Nsɛmfua yi bi wunim, na ebi nso wunnim wɔ Twi kasa mu . Nsɛmfua yi a wunim no ɛtɔ asom, na nea wunnim no ntɔ asom wɔ Twi kasa mu . Wobɛkenkan ne nyinaa. Sɛ asɛmfua no wunim a, san (√) wɔ adaka no mu. Sɛ asɛmfua no wunnim a, fa (X) hyɛ adaka no mu. Ma yɛnyɛ nhwɛso ahorow yi.**

[Write the word 'kew' with a box beside it on the board]. **Ma yɛnyɛ nhwɛso. Asɛmfua a ɛdi kan no yɛ 'kew'.** [Point to 'kew'.] **Asɛmfua yi wunnim wɔ Twi kasa mu. ɛno nti mede (x) bɛhyɛ adaka no mu.** [Put a wrong mark X in the box.] **Fa wo nsateaa si asɛmfua 'kew'.** [Check to see if they are pointing to 'kew'.] **Afei fa (X) hyɛ 'kew' adaka no mu.** [Check to see if they put an x mark next to 'kew'.]

[Write the word 'dua' with a box beside it on the board.] **Ma yɛnyɛ nhwɛso foforo. Fa wo nsateaa si asɛmfua ' dua ' wɔ wo nhoma no so na kyerɛ sɛ wunim asɛmfua yi wɔ Twi kasa mu. Afei fa (✓) hyɛ ' dua ' adaka no mu.** [Check that they all marked an answer.]

Afei wo ankasa wobɛyɛ eyinom. Sɛ meka sɛ fi ase a, fi ase wɔ nɔma 1 kosi 30. Fa wo nsateaa si adaka 1 no so. [Check to see if they are in the correct spot] **Kenkan asɛmfua biara, na hwɛ sɛ wunim asɛmfua no anaasɛ wunnim. Dɛn na yɛde ma asɛmfua a wunim no?** [Give children a chance to respond. Correct as needed.] **Afei dɛn na yɛde ma asɛmfua a wunnim no?** [Give children a chance to respond. Correct as needed.] **Yɛwɔ simma abiɛsa ma dwumadi yi. Wo yɛ krado. Fi ase.** [Start timer. Monitor. At 3 min., stop the task.]

Word		Child response*		Word		Child response*	
1	bɔ	*I know*	*I don't know*	16	ɔdra	*I know*	*I don't know*
2	ɔbra	*I know*	*I don't know*	17	haro	*I know*	*I don't know*
3	li	*I know*	*I don't know*	18	sukuu	*I know*	*I don't know*
4	didi	*I know*	*I don't know*	19	mpa	*I know*	*I don't know*
5	rɔ	*I know*	*I don't know*	20	ubɛ	*I know*	*I don't know*
6	ɔpra	*I know*	*I don't know*	21	rek	*I know*	*I don't know*
7	nɔ	*I know*	*I don't know*	22	ɔdan	*I know*	*I don't know*
8	rem	*I know*	*I don't know*	23	edin	*I know*	*I don't know*
9	popa	*I know*	*I don't know*	24	sek	*I know*	*I don't know*
10	mam	*I know*	*I don't know*	25	wɔfa	*I know*	*I don't know*
11	nak	*I know*	*I don't know*	26	kwe	*I know*	*I don't know*
12	mfe	*I know*	*I don't know*	27	nifa	*I know*	*I don't know*
13	kɛse	*I know*	*I don't know*	28	hese	*I know*	*I don't know*
14	asia	*I know*	*I don't know*	29	ban	*I know*	*I don't know*
15	goma	*I know*	*I don't know*	30	gɔm	*I know*	*I don't know*

*On this protocol, do not mark if the child was correct or incorrect. Just note child's response

Mo, woayɛ ade. Ma yɛnkɔ ɔfa a edi so no. Good effort! Let's go on to the next page.

Group of 4-6 children.

Start the timer after completing the examples

🖐 When the timer reaches 0, say "stop."

Subtask: Sentence Choice	🕮 Children write in booklet.	🕐 5 minutes

| | | Group of 4-6 children.

Start the timer after completing the examples.

✋ When the timer reaches 0, say "stop". |

🕪 **Fa wo pɛnsere no to hɔ na tie me.** [Write a large √ and an **X** on the board. Ask the children what they call these symbols.]. **ɔfa ha yɛwɔ kasamu ahorow bi wɔ Twi kasa mu.** [Point to the page with sentences.] **ɔkasamu no bi** <u>**yɛ nokware**</u> **na bi nso** <u>**nyɛ nokware**</u>**. Wo ankasa bɛkenkan ne nyinaa. Sɛ kasamu no** <u>**yɛ nokware**</u> **a, san (√) ma no. Sɛ kasamu no** <u>**nyɛ nokware**</u> **a, san ho (X). Ma yɛnyɛ nhwɛso ahorow bi.**

[Write the sentence 'Akokɔ wɔw fufuu' with a box beside it on the board.] **Nhwɛso a edi kan: 'Akokɔ wɔw fufuu'.** [Point to the sentence.] **So Akokɔ wɔw fufuu? Daabi. Eyi nyɛ nokware efisɛ Akokɔ nwɔw fufuu. Yɛde (X) bɛhyɛ adaka no mu.** [Write an X mark in the box.] **Fa wo nsateaa si ɔkasamu 'Akokɔ wɔw fufuu'.** [Check to see that they are in the correct spot.] **Afei fa (X) hyɛ 'Akokɔ wɔw fufuu' adaka no mu.** [Check to see if they put an x mark next to the sentence.]

[Write the sentence 'Agyinamoa we akura' with a box beside it on the board.] **Ma yɛnyɛ foforo bio. 'Agyinamoa we akura'.** [Point to the sentence.] **So Agyinamoa we akura? Yiw. ɔkasamu yi yɛ nokware, efisɛ agyinamoa we akura. Ɛno nti yɛde (√) ma no.** [Write a tick (√) in the box.] **Fa wo nsateaa si ɔkasamu 'Agyinamoa we akura' no so.** [Check to see that they are in the correct spot.] **Afei fa (√) hyɛ 'Agyinamoa we akura' adaka no mu.** [Check to see if they put a tick (√) next to the sentence].

Afei wo ankasa wobɛyɛ eyinom. Sɛ meka sɛ fi ase a, fi ase wɔ nɔma 1 kosi 20. Fa wo nsateaa si adaka nɔma 1 no so. [Check to see if they are in the correct spot] **Kenkan ɔkasamu no na hwɛ sɛ ɛyɛ nokware anaasɛ nyɛ nokware. Sɛ ɔkasamu no yɛ nokware a, dɛn na yɛde ma no?** [Give children a chance to respond. Correct as needed.] **Sɛ ɔkasamu no nyɛ nokware a, dɛn na yɛde ma no?** [Give children a chance to respond. Correct as needed.] **Sɛ meka sɛ gyae a, fa wo pɛnsere no to hɔ. Yɛwɔ simma anum ma dwumadi yi. Wo ayɛ krado. Fi ase.**
Start timer. Monitor children. At 5 minutes, stop the task.

Sentence		Child response*		Sentence		Child response*	
1	Dua tow nkesua.	True	False	11	Yɛde yɛn nan kasa.	True	False
2	Yɛde awia guare	True	False	12	Ɛsono sua.	True	False
3	Yedi ɔbo.	True	False	13	Yedi kwadu.	True	False
4	Ɔkraman te nsu mu.	True	False	14	Dɔte yɛ dɛ.	True	False
5	Atɛtea sua.	True	False	15	Nnipa ka lɔre.	True	False
6	Anomaa ka lɔre.	True	False	16	Mmoa kɔ sukuu	True	False
7	Akokɔ tow nkesua.	True	False	17	Tɔfe yɛ dɛ.	True	False
8	Mmofra kɔ sukuu.	True	False	18	Yɛde nsu guare.	True	False
9	Ɔkɔtɔ te nsu mu.	True	False	19	Oguan kan nhoma	True	False
10	Sukuufo kan nhoma	True	False	20	Yɛde yɛn ano kasa.	True	False

*On this protocol, do not mark if the child was correct or incorrect. Just note child's response

Mo woayɛ ade. Ma yɛnkɔ ɔfa a edi so no. Good effort! Let's go on to the next page.

Subtask: Word Dictation	📖 Children need ruled paper	🕐 4 minutes

🎤 Fa wo pɛnsere no to hɔ na tie me. ɔfa ha yɛwɔ Laen dumien (1–12) wɔ nhoma yi so. [Point to the paper.]Yɛbɛkyerɛw Twi nsɛmfua bi wɔ laense no so. Mɛyɛ nea edi kan no. Mepɛ sɛ mekyerɛw asɛmfua 'saw'. ɛno nti merekyerɛw 'saw' wɔ laen a edi kan no so. [Write the word 'saw' on the board as you say it.]

Afei wobɛkyerɛw nsɛmfua bi. Sɛ woantumi ankyerɛw nsɛmfua no mu nkyerɛwde no nyinaa koraa a, ɛnyɛ hwee. Kyerɛw nkyerɛwde a wote wɔ asɛmfua no mu no.

Yɛnyɛ nhwɛso bi. Fa wo nsateaa si laen a edi kan no so. [Check to see that all the children are on the line labeled Example.] Kyerɛw asɛmfua ' bra'. [Point to the left top line.] Kyerɛw asɛmfua no mu nkyerɛwde no. [Wait for the children to write the word 'bra' on the first line. If a child does not start, say, **Just write the letters that you know in the word.**] Sɛ wunnim a, ɛnyɛ hwee.

[After the children have attempted the practice word say, I can tell you are trying.] **Mo, woayɛ ade. Sɛ meka 'fi ase' a, fi ase wɔ nɔma 1 na kyerɛw asɛmfua a wote no.** [Check to see that all the children are on number 1]. Sɛ wunwiee asɛmfua no na meka sɛ toa so a, kɔ laen a edi so no na kyerɛw asɛmfua a meka no wɔ hɔ. Yɛwɔ seconds aduonu ma dwumadi yi. Sɛ meka sɛ gyae a, fa wo pɛnsere no to hɔ. [Read each word. Repeat at 3 seconds. At 10 seconds say the last word a third time. Stop the activity after 4 minutes].

Right column: Group of 4-6 children.

Start the timer after completing the example

🖐 When the timer reaches 0, say "stop."

.

Time	Word	Put a tick* in each box to note how the child spelled the word.						
:00	**fa**	_f	fa	a_	Word correct			
:20	**da**	_d	da	a_	Word correct			
:40	**bu**	_b	bu	u_	Word correct			
:60	**abɛ**	_a	ab	bɛ	ɛ_	Word correct		
1:20	**ɔbo**	_ɔ	ɔb	Bo	o_	Word correct		
1:40	**kan**	_k	ka	An	n_	Word correct		
2:00	**tɔn**	_t	tɔ	ɔn	n_	Word correct		
2:20	**pam**	_p	pa	Am	m_	Word correct		
2:40	**soma**	_s	so	Om	ma	a_	Word correct	
3:00	**horo**	_h	ho	Or	ro	o_	Word correct	
3:20	**ɔsra**	_ɔ	ɔs	Sr	ra	a_	Word correct	
3:40	**kyerɛ**	_k	ky	Ye	er	re	e_	Word correct

*Cross out the boxes that do not match the way the child wrote the word. If the entire word is correct, circle word correct.

Mo! Woayɛ ade. Yɛnkɔ ɔfa a edi so no so. Good effort! Let's go on to the next page.

Subtask: Semantic Fluency	▤Assessor writes in blank page in pupil booklet	⏲ 60 seconds

		Start the timer after explaining the task.

🕮 Nea edi kan yebedi nkɔmmɔ kakra. Dɛn na wopɛ sɛ woka fa wo nnamfonom ne w'abusuafo ho? [Wait for response.] For this activity we will just talk. What do you like to talk about with your friends and family?

Subtask 1a. **Mepɛ sɛ me kasa fa me honam akwaa ho. Me nsa ni.** [Show your hand.] **Mepɛ sɛ wobobɔ wo ankasa wo honam akwaa no din dodow biara a wunim wɔ Twi kasa mu kyerɛ me, kosi sɛ mɛka sɛ eye. Wutumi ka honam akwaa biara a wunim na makyerɛw. Moate ase?** [If the student says no, repeat instruction]. **Fi ase.**

Start the timer. Write down all the parts of body the child says.
 After 60 seconds say: **Gyae. Mo. Woayɛ ade. Ma yɛnka ade foforo bi ho asɛm.** Stop. I can tell you are trying. Let's do one more talking activity.

 Subtask 1b. **Yɛaka honam akwaa ho asɛm. Afei mepɛ sɛ wobobɔ nsɛmfua biara a wunim wɔ Twi kasa mu kyerɛ me, kosi sɛ mɛka sɛ eye.   Wutumi ka asɛmfua biara a wunim wɔ Twi kasa mu na makyerɛw. Moate ase**? [If the student says no, repeat instruction]. **Fi ase.**

Start the timer.   Write down all of the words that the child says.
After 60 seconds say: **Gyae. Mo. Woayɛ ade.** Stop. I can tell that you are trying.

Scoring note to assessor: After the child completes the task, ensure that you can read what you have written. For each topic, use the acronym LORU. 1) Write the letter L next to words said in another language. 2) Write the letter O next to words that are off topic. 3) Write the letter R next to words that are repeated. 4) Total the L words and then cross them off. 5) Total the O words and then cross them off. 6) Total the R words and then cross them off. 7) The remaining words will be recorded as the total unique words (U) for that category.

| | ✋ When the timer reaches 0, say "stop".

If the child gives you a response in a language other than Akuapem Twi, say **Please tell me in Akuapem Twi**. This prompt may be given once per category during the subtask. |

Subtask 1a (Animals)	Subtask 1b (Body parts)	Subtask 1c (Any category)	
1. Total words in another language |___|___|___|	1.Total words in another language |___|___|___|	1. Total words in another language |___|___|___|	
2 Total words off topic |___|___|___|	2 Total words off topic |___|___|___|	2. Total words off topic Off Topic is not relevant and not scored for Any Category	
3. Total repetitions |___|___|___|	3. Total repetitions |___|___|___|	3. Total repetitions |___|___|___|	
4. Total unique words |___|___|___|	4. Total unique words |___|___|___|	4. Total unique words |___|___|___|	

Mo, woayɛ ade.