**What is the EGRA, and why create new 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 (see RTI International, 2016a). As of early 2017, EGRA had been adapted into over 100 languages for use by more than 50 organizations in over 70 countries. It is administered one-on-one between a trained assessor and an individual, usually a student in primary school.

During the 10 years of its application, the EGRA has been a useful tool to understand students’ progress toward fluent reading. However, users are often left wanting additional information about reading comprehension, writing, and language.

The subtasks that we piloted in the research activity described in this brief can give researchers and practitioners more detailed information to understand the early literacy abilities of students in low-income contexts.

The EGRA has been a useful tool to understand students’ progress toward fluent reading. However, users are often left wanting additional information about reading comprehension, writing, and language.

The subtasks that we piloted in the research activity described in this brief can give researchers and practitioners more detailed information to understand the early literacy abilities of students in low-income contexts.

During the 10 years of its application, the EGRA has been a useful tool to understand students’ progress toward fluent reading. However, users have also noted their desire for additional, more nuanced information about other basic literacy skills. This brief describes a 2016 pilot study carried out by RTI International—with funding from USAID—to better understand students’ reading comprehension, writing, and oral language skills (RTI International, 2016b). It involved the development and piloting of four additional subtasks.
Skills assessed

Spelling

Although the existing EGRA has a Dictation subtask that measures spelling, its drawback is binary scoring, which limits what can be learned about the development of early spelling skills. The new subtask involves a more qualitative scoring approach.

SUBTASK: **Word Dictation.** This subtask captures the developmental nature of spelling skills using a scoring protocol awarding partial correctness. Twelve words are scored for correct letter sequences, or pairs of letters written in the proper sequence. This yields an overall correct/incorrect score for each word, as well as a more nuanced score equal to the number of letters plus one. For example, the word “talk” was scored for five items:

- Sequence 1: space t (_t)
- Sequence 2: t a (ta)
- Sequence 3: a l (al)
- Sequence 4: l k (lk)
- Sequence 5: k space (k_) (i.e., _t ta al lk k_)

Word Dictation can be administered individually or to a small group of students.

Reading comprehension

Among the concerns with EGRA’s existing Reading Comprehension measure is the fact that the number of comprehension questions administered is determined by the number of words in the passage that the student attempts to read in one minute. Thus, one goal was to be able to expose students to more items to better understand their reading comprehension abilities. We piloted two additional reading comprehension subtasks, using a protocol in which all items are attempted, regardless of reading rate.

SUBTASK: **Word Choice.** This subtask measures word-level comprehension by mixing 15 individual real words from the language of assessment with 15 orthographically legal nonwords. To reduce measurement error, the student evaluates each word as either “real” or “fake” (e.g., “tick/check [✓] the real words and put an X next to fake words”). Word Choice can be administered individually or to a small group of students.

SUBTASK: **Sentence Choice.** This subtask measures comprehension of 20 independent sentences. Based on their everyday knowledge, students discern whether the sentence is either obviously true or obviously false (e.g., “Goats fly in the sky”). Sentence Choice can be administered individually or to a small group of students.

Oral language

The existing Listening Comprehension and Vocabulary subtasks of the EGRA both successfully measure receptive language skills. The pilot’s intent was to better gauge learners’ expressive oral language skills and attention.

SUBTASK: **Semantic Fluency.** This subtask measures verbal ability, lexical knowledge, and lexical retrieval. Students are first asked to retrieve as many unique words as they can in one minute within a provided semantic category (e.g., animals). Next, they are asked to say as many unique words as they can think of (in one minute), from any category (i.e., any unique word in the language being assessed). Semantic Fluency is administered individually.

Pilot

In mid-2016, we oversaw initial adaptation, refinement, and multiple field tests of the four proposed subtasks in Accra, Ghana. Next, we piloted them with students in lower primary, in English and in Akuapem Twi. The pilot administration also included three existing EGRA subtasks (Nonwords, Oral Reading Fluency, and Reading Comprehension) to explore the relationships between piloted and existing subtasks and to determine possible redundancy.

Technical Adequacy

We conducted reliability tests on all of the piloted subtasks to determine their internal consistency (reliability) and to gauge whether they were measuring the intended construct (validity). The findings from the reliability tests are summarized below.

Word Dictation. Our analysis showed that the piloted scoring methodology yielded variability in results. Most useful was that even for students who did not spell any of the words correctly, they could correctly represent the first letter, last letter, or a salient sound, which parallels how spelling develops. Our analysis showed this to be a reliable measure for understanding writing skills in lower primary.
Word Choice. Our analysis showed that Word Choice had good internal consistency (0.80) but that it had weak construct validity and test-retest reliability. Its relationship (0.54) to Nonwords suggested it was not measuring the decoding construct as intended. Nor did it have a strong relationship to reading with understanding (0.55 to Sentence Choice). The unreliability of this subtask was also seen in the test-retest comparison (0.69) that we conducted on the same day, in which students changed their answers, thereby benefiting from a practice effect. This subtask is not recommended.

Sentence Choice. Our analysis showed Sentence Choice to be a reliable comprehension measure for use in for lower primary. The test scale for paired sentence scores showed strong signs of internal consistency (0.80 in the initial testing and 0.74 in the final pilot). Item analyses functioned as expected and the relationship between scores at two time points was adequate (0.76). Our analysis suggested that the percentage of students scoring correctly was beyond that of chance alone and was related to their results on the Oral Reading Fluency subtask. Conversely, students unable to read a single word on Oral Reading Fluency showed response patterns on Sentence Choice that suggested they were guessing. Also, Sentence Choice’s relationships with the existing EGRA Oral Reading Fluency (0.68) and Reading Comprehension (0.67) subtasks were strong enough to suggest this simplified measure requires both reading accuracy and meaning.

Semantic Fluency. Our analysis of the results from this new subtask were inconclusive. Students seemed to enjoy this activity and the results were reliable. However, to further explore the constructs, it would be beneficial to use it alongside the existing EGRA receptive language subtasks and an expressive language measure containing images (e.g., Prompt: What is this? Answer: A bird.)

Recommendations
The pilot testing led us to make the following recommendations about the additional EGRA subtasks.

- Use Sentence Choice as a complement to or replacement for the existing Reading Comprehension subtask. Children in lower primary in Ghana were able to understand the procedures of this subtask. The English version had a strong relationship to the existing Reading Comprehension measures, and the scores were stable. However, the Akuapem Twi version did not perform well; we recommend further development of the sentence structure guidelines.
- Conduct additional piloting with Semantic Fluency to examine its relationship to existing language measures. It was appealing to students and its simplicity to adapt seemed promising.

References


More Information
Margaret (Peggy) Dubeck
+1 202 728 2056
dubeck@rti.org
Follow @RTI_EdWork
Follow @RTI_Intl_Dev
http://ShareEd.rti.org
RTI International
3040 E. Cornwallis Road, PO Box 12194
Research Triangle Park, NC 27709-2194 USA