



USAID
FROM THE AMERICAN PEOPLE

EGRA Assessor Recruitment, Evaluation, and Selection

Prepared for the USAID workshop
“Designing and Implementing Early Grade Reading
Assessments: Understanding the Basics”

March 2015

Recap of Where We Are

- Identification of Research Design and Sampling Framework
- Development/Adaptation of EGRA Instrument
- Procedures for EGRA Administration, Scoring and Data Capture
- Establishment of Electronic Data Capture System
- **Enumerator Training, Assessment and Selection**
- Pilot and Full Data Collection
- Use and Dissemination of EGRA Results
- Planning and Managing EGRA Implementation

Session Objectives

- Be knowledgeable about best practices, strategies, and tools for monitoring and evaluating data collector performance
- Understand the purpose of interrater reliability testing, methods for conducting, and how to use and interpret results

Recruitment and Selection of Assessors and Supervisors

- Data collection teams may be composed of education officials and/or independent enumerators recruited for a particular data collection.
- Requirements and preferences should be determined during the recruitment phase, in advance of the training, depending on specific circumstances.

Assessor Recruitment and Selection Criteria

An assessor should meet the following criteria before being selected:

- Fluently read and speak the language being assessed and the language in which instructions are provided to pupils.
- Have experience working with young children.
- Be able to follow directions and pay attention to detail.
- Have experience using tablet device, computer, or android phone.
- Have previous experience participating in data collection.
- Work well as part of a team.
- Be available for the entire duration of data collection.
- Be able to travel to rural and remote areas.
- Be employed by a relevant education institution (if applicable).
- **Demonstrate ability to accurately administer EGRA during the training.**

Supervisor Criteria

A supervisor must meet all of the assessor criteria, plus:

- Have experience effectively leading a team.
- Be organized.
- Know EGRA administration procedures well enough to supervise others and check for mistakes in data collection.
- Possess sufficient knowledge/skills of tablet devices in order to help others.
- Exhibit the ability to interact well with school officials and children.

Government vs. Nongovernment Data Collectors

- Data collection teams can be composed of education officials and/or independent enumerators.
- Requirements and preferences should be determined during the recruitment phase, in advance of the training, depending on specific circumstances.

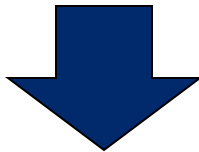
Factors to consider:

- Sustainability
- Capacity development
- Acceptance of results
- Regulations regarding government participation
- Availability
- Potential conflict of interest

Selecting Assessors and Supervisors

- Communicate the process and criteria for assessor selection at the beginning of the workshop.
- Selection should be based on evaluation of performance during the workshop.
- In addition to being competent assessors, supervisors should exhibit leadership and organizational skills.

Data Collection Team: Indonesia's case

- Nonteachers and nongovernment officials
 - Time commitment: 10 days training, 5 days piloting, 30 days data collection
 - Familiarity with gadget/electronic devices
 - Young, strong, and healthy physical condition to complete the process
- 
- University students: Faculty of Education, majoring in Bahasa Indonesia language
 - Above-average academic achievement, 7th semester and above
 - Carefully selected prior to the training via interviews
 - 1 team: 2 female and 2 male assessors, for efficiency purposes
 - All supervisors were male

Assessor Teams: Indonesia's Case

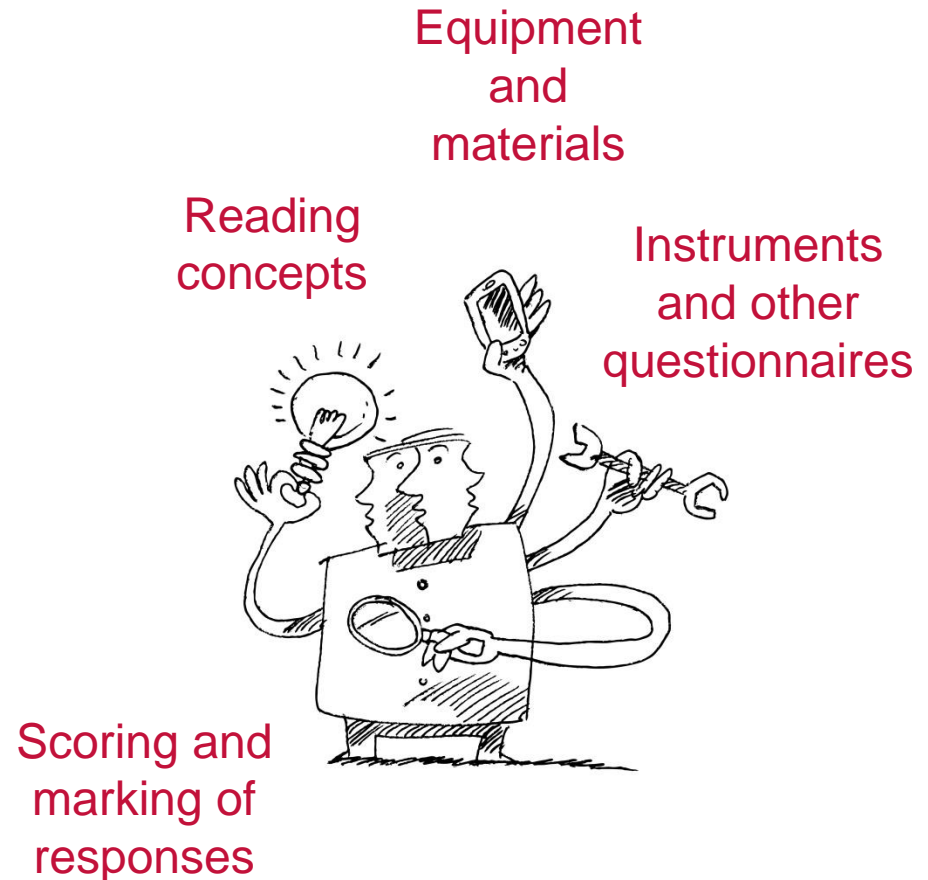
- All assessors signed an individual agreement before the training to abide by their commitment.
- All assessors were insured (life and health insurance) to protect them during the piloting and data collection process.
- Assessors were assigned as a team, so that if one of them was unable to fulfill duties during the process (due to illness or inability), all of the team members were replaced with another team.

Assessor Teams: Indonesia's case



Why Evaluate Assessor Performance?

- Tailor training to trainees' specific needs
- Identify trainees in need of additional assistance
- Select the best trainees to serve as assessors for data collection
- Gather accurate, high-quality data
- Provide quantifiable information about the quality of the data captured and analyzed



Areas for Assessor-Trainee Evaluation

- Evaluation of assessor-trainees is multifaceted and takes into consideration several factors, including the ability to:
 - Correctly and efficiently administer instruments
 - Accurately record demographic data and responses
 - Identify responses as correct and incorrect
 - Correctly and efficiently use equipment, especially tablets
 - Work well as a part of a team
 - Adhere to school visit protocols
 - Create a rapport with pupils and school personnel

Evaluating Assessor Performance

A transparent evaluation process and clear criteria for evaluation are helpful for both facilitators and trainees.

Methods

- Facilitator observations during workshop and school visits
- One-on-one interview
- Formal assessment

Tools

- Participant observation sheet
- Assessor observation checklist
- Test of interrater reliability (IRR)

Handout 8.1: Assessor Observation Checklist
Handout 8.2: Assessor Report Card
Handout 8.3: Participant Observation Sheet

Helpful Tip

Take a photograph of each participant at the workshop to help with trainee identification during the evaluation process. Pictures can be taken on the first day of the training when participants arrive and sign in. Ask participants to wear or hold up a name badge.

Handout 8.1: ASSESSOR OBSERVATION CHECKLIST

Session 8: Data Collector Assessment and Selection

PRIOR TO TEST ADMINISTRATION	Yes	No	Remarks
1. Assessor prepares the assessment space appropriately (no materials on the table/desk except the pupil stimuli) and all materials are ready when the child sits down. Switches off mobile.			
2. Assessor is relaxed, makes the child feel comfortable and establishes a good rapport.			
3. Assessor reads the consent verbatim. In the case where the child does not wish to participate, the assessor thanks the child and gives him/her the small gift.			
ADMINISTRATION OF SPECIFIC SUB-TASKS	Yes	No	Remarks
Sub-task 1: Letter Sounds			
• Assessor turns to correct page in the stimulus book and places it before the child before beginning to read the directions.			
• If the learner hesitates for 3 seconds before attempting the next item, assessor says “please go on” and points to the next letter.			
• If learner gives letter name, assessor says “tell me the sound of the letter”. Prompt IS given only once.			
• Assessor discontinues the sub-task if the child’s responses to the first 10 letters (first line) are incorrect.			
Sub-task 2: Non-word Reading			
• Assessor turns to correct page in the stimulus book and places it before the child before beginning to read the directions.			
• If the learner hesitates for 3 seconds before attempting the next item, assessor says “please go on” and points to the next word.			
• Assessor stops task if no words correct in first line/first 5 words.			

Handout 8.2: Assessor Report Card

Session 8: Data Collector Assessment and Selection

Name	XXXXXXXXXX
Language	<u>Akuapem Twi</u>
Username	XXXXX

Tablet Care: Returned in Good Condition

General Tangerine Use

Used One Username	No
Used Dropdown Menu for School Selection	Yes

Subtask Assessor Performance

Assessor Action	EGRA English	EGRA <u>Akuapem Twi</u>	EGMA
Assessor stopped the time when pupil completed all items in the subtask	Well Done	Well Done	Well Done
Assessor stopped the time ONLY when the pupil completed all items in the subtask	Well Done	Well Done	Well Done
Assessor nudged the pupil to move on after the correct number of seconds, and neither rushed the pupil nor allowed the pupil to try for too long on each item	Needs Improvement (too slow)	Well Done	Poor (too fast)

Handout 8.3: PARTICIPANT OBSERVATION SHEET

Session 8: Data Collector Assessment and Selection

During the training, use this sheet to record participant progress. Rate each participant in the four categories. Conduct a one-on-one interview with each participant by the end of the training. Ask the participant to conduct the Hausa assessment through letter sounds. Rate the participant in each category as follows:

Excellent: ✓+ Good (room to improve): ✓ Needs significant improvement: ✓-

	Participant Name	<u>Reading instructions</u> <ul style="list-style-type: none"> • reads verbatim • reads clearly • reads at a good pace (slow enough for participant to understand, yet fast enough to efficiently move through instrument) 	<u>Marking responses</u> <ul style="list-style-type: none"> • knows correct letter sounds (EGRA) • records participant responses accurately • respects protocol for marking items as incorrect, correct, etc. • marks efficiently (does not have to try multiple times to select/tap) • keeps "perky" pace (does not delay) 	<u>Tablet use</u> <ul style="list-style-type: none"> • able to find and start instrument • scrolling – easily moves from top to bottom • keeps pace with pupil responses on EGRA and participant responses on questionnaires • knows how and when to "stop" the timer in Tangerine if applicable • moves quickly from one section to the next • able to resume an instrument • able to save and upload data 	<u>Overall administration, comportment</u> <ul style="list-style-type: none"> • establishes good rapport • knows when to give pupil stimuli (EGRA) or response sheet (KSA) • knows rules for auto-stop, 3- and 5-second rules, etc. in EGRA • knows proper sampling procedures, set-up at schools • works well with other assessors
1.					
2.					

Interrater Reliability

- **Definition: Interrater reliability (IRR)** is the degree of agreement among multiple raters (i.e., assessors) judging the same test.
- **Purpose:** To evaluate an assessor's ability to consistently and accurately assess and record pupils' responses on EGRA
 - For example, the degree to which an assessor scores the items correctly (i.e., marks a “b” wrong if the child pronounced it incorrectly, and vice versa)
- **Frequency:** Conducted approximately 3 times during the course of the training
- **How it's used:** Informs training (focus on specific problematic items), aids in the selection of assessors, and supports the reliability of EGRA results and data analysis

Steps for Conducting IRR

- Assessor-trainees are assembled into one large group.
- A mock assessment of a child is given, using either two adult trainers (with one posing as a child), a video, or a combination of the two.
- The child's responses are based on a script prepared beforehand—known as the **Gold Standard**—so that the errors made are deliberate and unambiguous.
- The assessor-trainees record the “child's” responses as they would during EGRA administration. Their results are captured and used to calculate their IRR scores.

Conducting IRR testing

Trainers play the role of student and assessor



Participants mark their responses



Activity: Practice IRR

Calculating IRR Scores: Percentage Agreement

enumerator	Non word	non_word_time_remain	non_word_attempted	non_word1	non_word2	non_word3	non_word4	non_word5	non_word6	non_word7	non_word8	non_word9	non_word10	non_word11	non_word12	non_word13	non_word14	non_word15	non_word16	non_word17	non_word18	non_word19	non_word20	non_word21	non_word22
Average	88%																								
aloreirr1	91%	1	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
apanjirr1	81%	1	0	0	1	0	1	1	0	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1
ashooirr1	75%	1	0	0	1	0	1	1	0	1	1	1	1	1	0	1	0	1	0	1	1	1	1	0	0
dmtitirr1	89%	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
hseleirr1																									
ikiwairr1																									
jmasairr1																									
jurasirr1																									
kkahairr1																									
lkayoirr1																									
mkyejirr1	79%	1	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1	0	0	1	0	1	1	0	1
mndolirr1	93%	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
mpaziirr1	91%	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1
mramairr1	91%	1	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1

$$\text{IRR score (subtask/assessment) (\%)} = \frac{\text{Number of agreements with the Gold Standard}}{\text{Number of items in the subtask/assessment}}$$

Steps and Best Practices for Conducting IRR

Before

- Create “IRR instrument” in tablets
- Develop a “Gold Standard script” and practice in advance
- Harmonize across languages/training sites
- Prepare IRR testing space/logistics

During

- Use appropriate-language speakers
- Record Gold Standard student responses

After

- Review data
- Calculate and analyze results
- Provide results to assessors (group and individual)