



**USAID**  
FROM THE AMERICAN PEOPLE

# **Pilot and Full Data Collection**

## **Data Collection Preparation and Management: Key Tasks and Best Practices**

Prepared for the USAID workshop  
“EGRA Basics and Beyond”  
November 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

- Understand key tasks and best practices related to data collection preparation and management
- Review budgetary and logistics considerations, including equipment and supplies needed
- Understand scheduling constraints and requirements
- Understand the roles and responsibilities of various players

# Research Best Practices

- When conducting a pilot or full data collection, ensure the quality of the instruments being tested.
- When conducting fieldwork, ensure that you are following best research practices, including but not necessarily limited to:
  - Ensuring the safety and well-being of children being tested
  - Maintaining the integrity of your instruments
  - Ensuring that data are collected, managed, and reported responsibly (quality, confidentiality, and anonymity)
  - Rigorously following the research design

# Regulatory Frameworks Regarding Research and Data

- Federal Policy for Protection of Human Subjects (USA)
  - USAID ([22 CFR Part 225](#)) has signed onto this “[Common Rule](#)”
  - Requires review and assessment of risks and benefits by Institutional Review Board (IRB) in advance of survey
  - Informed consent by participating individuals
- United Nations’ [Fundamental Principle of Official Statistics](#):
  - “Individual data collected by statistical agencies for statistical compilation, whether or not they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.”
- Data obtained through research [defined](#) as:
  - “A systematic investigation including research development, testing, and evaluation designed to develop or contribute to generalizable knowledge.”

# Institutional Review Board

- An **Institutional Review Board (IRB)** is an administrative body formally designated to approve, monitor, and review biomedical and behavioral research involving humans.
- History of IRB and human subject protections began with the Nuremberg Military Tribunals, key in Tuskegee syphilis cases.
- US regulatory protections were established May 1974.
- USAID, in addition to many organizations (and some countries), requires prior review of and approval of all survey instruments by IRB.
- Link to country IRB list:

<http://www.hhs.gov/ohrp/international/index.html>

# Institutional Review Board

- Research involving the use of educational tests in a school setting is typically “exempt” (based on principle that it is not substantially different from what children should already be doing). However, an IRB must determine exemption.

[http://en.wikipedia.org/wiki/Institutional\\_review\\_board](http://en.wikipedia.org/wiki/Institutional_review_board)

- Even if the research is classified as exempt, IRB must make that determination and all survey questions will require preapproval, especially if they seek to collect information that may put students or teachers at risk.
- If your organization does not have an IRB, you can partner with an organization that does.
- Commercial IRBs are also an option:

<http://www.circare.org/info/commercialirb.htm>

# IRB Best Practice



**Best Practice:** Know and follow all IRB guidance when working with children and factor sufficient time into your work plan for IRB review and approval. Failure to obtain IRB approval in advance can result in your data being prohibited from use and dissemination.

# What could go wrong?

- In your group, make a list of 10 things that might cause a data collection (pilot or full data collection) *to go horribly wrong*.
- A few groups will be called on to share out

# Why Conduct a Pilot?

- A pilot test is typically a small-scale preliminary study conducted prior to a full-scale survey. Pilot studies are used to test the following:
  - The logistics of implementing the study (e.g., cost, time, efficient procedures, and potential complications)
  - The personnel (i.e., during pilot testing, assessors are able to make beginner mistakes and learn from them. Assessors who are making too many errors can be removed from the assessment team)
  - The EGRA instrument(s) and any accompanying questionnaires



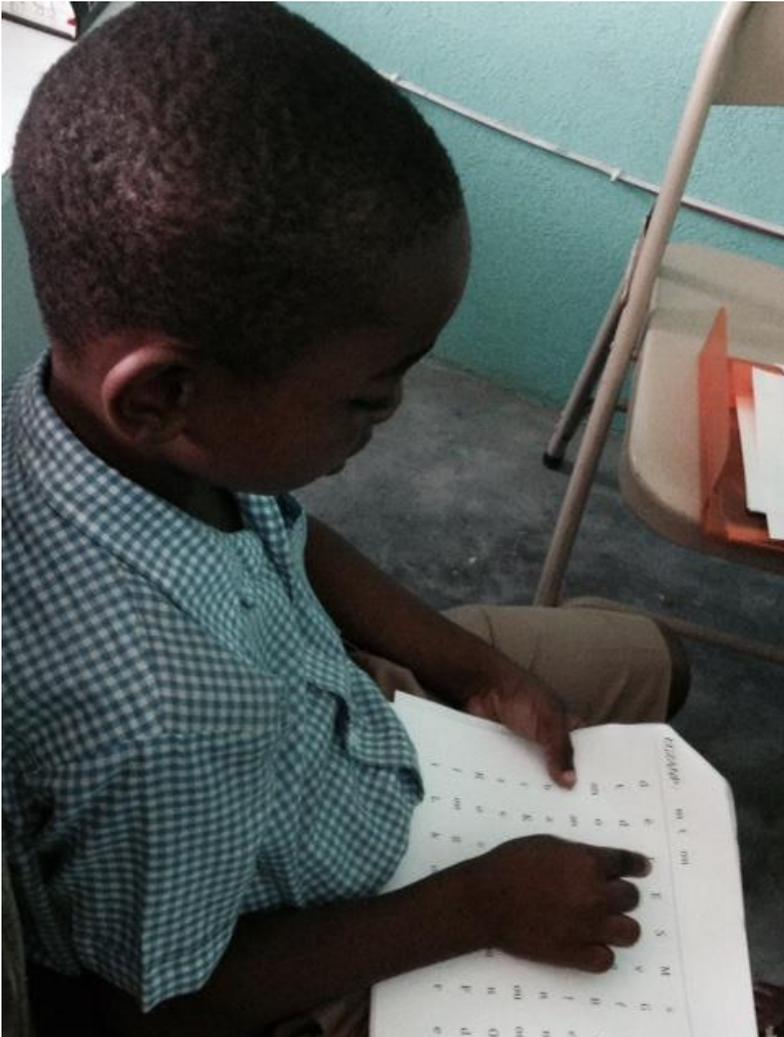
# Why Pilot Test Instruments?

- **Verify content of the EGRA instrument**
  - Content is appropriate for the target population (e.g., level of difficulty for the target grade level)
  - Items in each subtask are correct (no typographical errors, translation mistakes, or other errors)
  - New/unexpected responses to comprehension questions are documented
  - Pupil stimuli and assessor protocol content are the same
  - Wording of all instructions (for children and for assessors) is accurate, understandable, readable by the assessors, and error-free
  - Survey questions are culturally and age-appropriate, clearly worded, and error-free
  - Instrument can be administered in a reasonable and appropriate amount of time

# Why Pilot Test Instruments?

- **Review data collection procedures, logistics and materials**
  - Readiness of materials
  - Readiness and functionality of tablets and e-instruments
  - Feasibility of uploading data from the field
  - Transportation
  - Communication among assessor teams, field coordinators, and other project staff
- **Assess data collection personnel**
- **Obtain data to ensure instruments are valid**
  - Item-level responses
  - Equating

# Pilot Test Sample



- The sample and target population for an EGRA pilot is determined by several factors:
  - approximate skill level of children to be assessed
  - financial and human resources available
  - time available to conduct the pilot
- Pilot sample can be a convenience sample as long as the students being tested are comparable in key ways to those who will be assessed during the data collection (e.g., age/grade span, reading proficiency level, and socioeconomic status)

*See forthcoming EdData brief on pilot sampling for additional details*

# Testing Considerations

## Pilot Test

- Purpose
- Timing (e.g., after adaptation or after assessor training)
- Target population (grade, language, region, etc.)
- Content
- Planned analysis
- Revisions, if any, post-pilot

## Full Data Collection

- Purpose
- Timing (in relation to academic year, seasonal considerations)
- Target population (grade, language, region, etc.)
- Content
- Planned analysis
- Logistics

Pilot testing logistics should be as similar as possible to actual data collection, although not all subtasks may be tested and sampling considerations will likely vary.

# Pilot Test Best Practices

**Best Practice:** Develop an implementation plan for your full data collection in advance so that you can thoroughly test it during piloting.



**Best Practice:** Ensure that all key players are engaged, know the purpose, and understand the intended output of the pilot.



# Pilot and/or Full Data Collection Considerations

- Budget
- Equipment and supplies
- Support teams
- Supervision
- Scheduling and logistics
- In-school activities
  - Establishing rapport with school staff
  - Preparing the testing space
  - Sampling and selecting students

In both pilot study and full data collection, conducting frequent quality checks of the data being generated helps to identify problems and intervene as necessary.

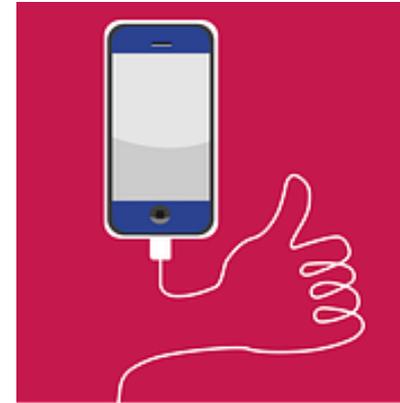
# Roles and Responsibilities

<b>Person(s)</b>	<b>Responsibilities</b>
<b>Subcontractor/ Local Project Staff</b>	<ul style="list-style-type: none"><li>• Arrange and oversee data collection logistics</li><li>• Select schools and students</li><li>• Supervise data collection teams</li><li>• Compile and send data to home office</li></ul>
<b>Assessor</b>	<ul style="list-style-type: none"><li>• Administer instruments per protocol</li><li>• Ensure all data are complete</li></ul>
<b>Supervisor</b>	<ul style="list-style-type: none"><li>• Oversee data collection, support enumerators as needed</li><li>• Ensure that protocols are followed and all data are complete</li><li>• Coordinate with subcontractor/local project staff, as needed, to replace schools</li></ul>
<b>Statistician/ Data Analyst</b>	<ul style="list-style-type: none"><li>• Review piloting plan</li><li>• Review data and conduct equating analyses</li><li>• Provide recommendations to team (including whether or not additional piloting is needed to attain enough nonzero scores)</li><li>• Daily review of data quality</li></ul>

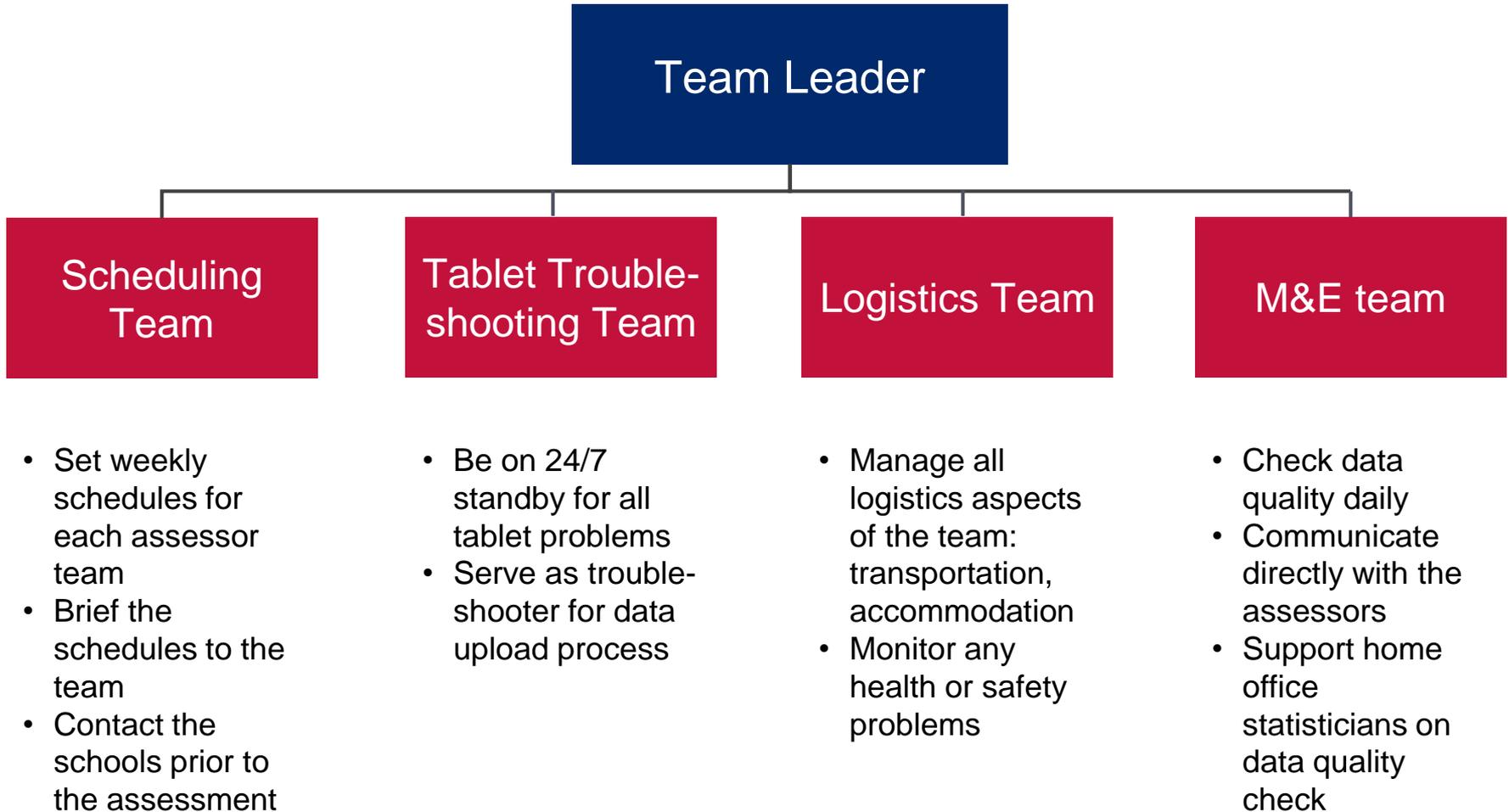
# Support Team

**Having a dedicated team on hand to support fieldwork will allow for immediate resolution of problems.**

- If using electronic tablets for data collection, have someone readily available to troubleshoot
  - power, functionality, and connectivity issues
  - the application being used.
- Assign someone to immediately respond to logistical and any other challenges and questions
- Assign a statistician or data analyst who can conduct regular (daily is optimal) data quality checks and provide feedback to the field as needed.



# Support Team Structure and Roles: Case Study



Supervisors and assessors focus only on data collection, resulting in good data quality

# Handout 11.1: Data Collection Assessor Agreement



## Handout 11.1: Data Collection Assessor Agreement Session 11: Pilot and Full Data Collection

The success of the Early Grade Reading Assessment (EGRA) to be conducted in COUNTRY/REGION depends on your participation and appropriate use of all material used during the data collection, which will take place from DATES. Your signature below signifies your agreement to fully participate in the data collection and to take care of all equipment that will be used during the course of the exercise. Specifically, you agree to:

- Arrive on time to the designated school every day.
- Respect the authority of the project staff, data collection colleagues, school personnel, and pupils.
- Use materials and equipment correctly and only as directed. In particular, this includes use of a tablet and all accessories (stylus, charger and cord).
- Use the tablet only as directed for data collection. The tablet should not be used for personal use at any time.
- Maintain the tablet and accessories in good working condition and keep them in a safe place at all times.
- Take care not to drop the device.
- Save and upload data as directed.
- Project a professional image during school visits; respect school property and leave classrooms as they are found.
- Work collaboratively with colleagues to efficiently accomplish tasks.

# Equipment and Supplies

- Depending on your research design, you will collect data either electronically or via paper instruments, and corresponding equipment will need to be acquired in or shipped to the appropriate location.
- Electronic
  - Sufficient time must be allowed to load and test tablets and copy student stimuli. Paper instruments are carried as a back-up method (so paper copies and stopwatches should also be provided).
  - Identify in advance how electronic devices will be charged, how data will be uploaded, and what backup batteries or wireless hotspots should be available.
- Paper
  - In large data collections, sufficient time for good-quality photocopying and transportation is a valid concern.
- Small gifts should be given to students after testing (e.g., pencil).

# Handout 11.2: Data Collection Supplies

## REVIEW THIS LIST BEFORE YOU GO TO EVERY SCHOOL!

The Team Supervisor is responsible for making sure that the team has all the materials needed each day.

<b>MATERIALS TO BE USED AT EACH SCHOOL</b> <i>Need to have a new supply every day.</i>	
<b>Number needed at each school per team</b>	<b>ITEM</b>
14	Small gift (pencil) for pupils (includes extras for alternates)
1	School visit authorization letter (to give to Head Teacher)
1	EGRA information brief (to give to Head Teacher)
1	<i>Pupil Sampling Worksheet</i> – completed at every school
1	<i>Teacher Sampling Protocol</i> – completed at every school
1	<i>School Fieldwork Summary Sheet</i> – completed at every school
<b>MATERIALS TO BE RE-USED</b> <i>Make sure they are in good condition and do not need to be replaced.</i>	
<b>Number needed at each school per team</b>	<b>ITEM</b>
1	List of sample schools to be visited by the team
1	List of codes for regions, schools, and assessors (in case paper data capture needed)
2/assessor	EGRA Pupil Stimuli
3	Tablets and accessories (1 per team member)
3	Clipboard (1 per team member)
2	Stopwatches (1 per team member)
1	Envelope to store back-up EGRA paper instruments
1	Envelope for Pupil Sampling Workshop and School Fieldwork Sampling Protocol
1	Marker to write on envelopes
1	Notepad for Supervisor

## PAPER BACK-UP INSTRUMENTS

# Supervision

- Supervisors are identified and trained during assessor training.
- One supervisor should accompany each team of assessors.
- Supervisors provide important oversight of assessors and should complete assessor rating forms for each school visit.
- Supervisors also support assessors by:
  - managing relationships with school staff
  - accompanying students to and from testing location
  - replenishing enumerators' supplies of testing documents or student gifts
  - communicating with the support team as needed
  - filling in as an assessor if needed

# Supervision Best Practices

## Best Practice:

- Plan for sufficient numbers of supervisors to allow one supervisor per school team.
- Select supervisors from assessor pool based on knowledge and demonstrated mastery.
- Meet with supervisors immediately prior to data collection to verify all logistics and distribute materials.



# Scheduling and Logistics

- Verify sample schools and identify location.
- Arrange travel logistics (transport and accommodations).
- Establish protocol for replacing schools, teachers, and/or students if needed.
- Develop an itinerary—critical to have a list of dates, schools, head teacher contact number, and team members.
- Organize and arrange delivery of assessment materials and equipment:
  - Backup copies of all final instruments
  - Relevant school sampling forms
  - Tablets and all equipment
  - School visit authorization letter.

# Scheduling and Logistics, continued

- Develop a clear plan and inventory system for maintaining and tracking equipment in consultation with local staff (use a tablet/equipment user agreement).
- Plan for data quality control checks during data collection (make sure statisticians are aware of timeline).
- Provide any additional training for supervisors, especially if there has been a long gap between training and data collection.

**Best Practice:** Be aware of the school calendar and identify any security, weather, or transportation concerns that may affect data collection—plan accordingly.

# Scheduling and Logistics Best Practice

**Best Practice:** Keep total data collection time to a minimum: 1–2 weeks is ideal.

- Reduces time for test to circulate
- Reduces impact of student growth on scores



# Transportation and Accommodation



# In-School Activities

- A group of people you don't know will arrive at your school and remove children from classrooms for testing. You a member of the community or are responsible in some way for the school. What questions do you have? Write your questions.
- Parent
- School director
- Teacher
- Student
- School inspector

# In-School Activities

- Team supervisor presents the team to the director—establishing a good rapport helps achieve better results and a more positive atmosphere for all involved.
- Explain the purpose of the visit: to conduct research regarding student reading achievement. The goal is to help the Ministry of Education improve education quality. Show the letter of introduction from the Ministry.
- **Emphasize** that the visit is **not** to evaluate the school, the principal, or the teachers. Note that all information will remain anonymous.
- Provide the principal with a 1-page brief on the assessment activity so they understand how results will be used.
- Answer any questions.

## In-School Activities, continued

- Explain the different activities that will take place (assessments, interviews, and classroom observations, as applicable).
- Tell the director that you will need to randomly select [#] pupils per grade, as well as some teachers (if applicable).
- Ask to use a classroom, or another quiet space, where assessors can administer the tests.
- Ask the principal to appoint someone at the school who can help the team throughout the day as needed. Identify a place where pupils can wait their turn.
- If applicable, set up a time to interview the principal and teachers and conduct classroom observations.

**Best Practice:** Obtain MOE approval and letters early, and follow all government protocols without divulging sample schools/testing dates.

# Handout 11.3: School Visit Summary Sheet

Date: \_\_\_\_\_

Supervisor:

--

Assessor 1:

--

Assessor 2:

--

Region:

--

District:

--

Name of School:

--

School Code:

--

School Tel. No.:

--

Head Teacher:

--

Arrival Time to School: \_\_\_\_\_:\_\_\_\_\_

Departure Time from School \_\_\_\_\_:\_\_\_\_\_

Will the school need to be revisited?  YES

NO

If YES, what day / date is planned to revisit: \_\_\_\_\_ (day)

Date: \_\_\_\_ / \_\_\_\_  
DD MM

Instruments administered

Primary 2 EGRA	
Primary 3 EGRA	
[specify language and any other instruments as applicable]	

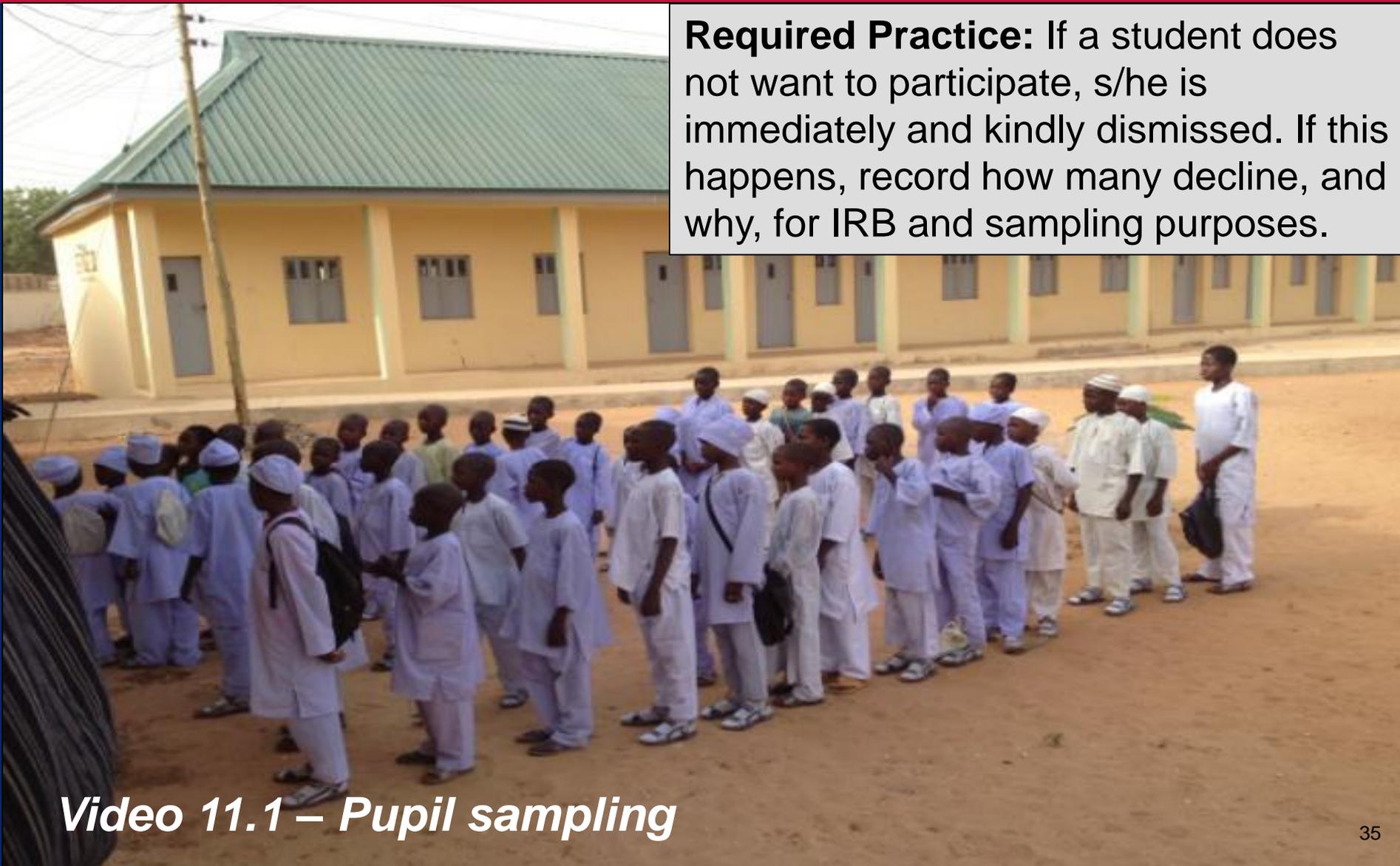
# Preparing the Testing Space

- The enumerator/supervisor team should work together to:
  - Locate a quiet space
  - Make sure students have sufficient light to read and view tablet screens
  - Arrange desks, and place materials in a way that reduces potential for student distraction
  - Silence mobile phones



# Pupil Sampling

**Required Practice:** If a student does not want to participate, s/he is immediately and kindly dismissed. If this happens, record how many decline, and why, for IRB and sampling purposes.



*Video 11.1 – Pupil sampling*

# Sampling Students

- Once the testing space is set up, supervisors, working together with assessors, should then randomly select the pupils to assess.
- At least two methodologies have been used for EGRA testing.
  - One involves identifying a “sampling interval” and selecting students accordingly
  - Another involves selecting students based on a random number table (which needs to be created by a statistician)
  - Either method is acceptable

# Handout 11.4: Pupil Sampling Worksheet – Sampling Interval

SCHOOL NAME: \_\_\_\_\_ REGION: \_\_\_\_\_ DISTRICT: \_\_\_\_\_ Date: \_\_\_\_\_

Class & Gender	Total number of pupils (Count boys and girls present in the class shift on day of sampling)			Sampling Interval (Use this number to identify pupils to test. Consult the Sampling Interval Calculations document for help with calculation if necessary.)
Primary 2		Divided by [Number of students needed]	=	
Primary 3		Divided by [Number of students needed]	=	

<b>Primary 2</b>					
	First Name	Surname	Section	Birth month	Birth year
1.					
2.					
3.					
4.					
5.					
6.					
7.					
8.					
9.					
10.					
11.					
12.					
13. (Alt. 1)					
14. (Alt. 2)					

# EXAMPLE: Pupil Sampling Worksheet – Sampling Interval

- To identify the sampling interval, divide the number of pupils present on the day of the assessment by the number of children needed; round appropriately

SCHOOL NAME: George Washington School STATE: Maryland DISTRICT: Bethesda Date: March 4, 2015

Class & Gender	Total number of pupils (Count boys and girls present in the class shift on day of sampling)			Sampling Interval (Use this number to identify pupils to test. Consult the Sampling Interval Calculations document for help with calculation if necessary.)
Primary 2	<b>51</b>	Divided by 12	=	<b>4.3 → 4</b>
Primary 3	<b>46</b>	Divided by 12	=	<b>3.8 → 4</b>

# Applying the Sampling Interval

- Based on a sampling interval of 4, we select every 4th pupil from the line of students/roster of pupils present

1. Muhammed

2. Safiatou

3. Lamin

4. Abdul

5. Aminata

6. Fatimata

7. Bilyaminu

8. Salisu

9. Abdoulaye

10. Mustafa

11. Binta

12. Ali

13. Abdul Rahim

14. Hawa

15. Fati

16. Aysata

17. Etc.....

# Sampling Students

- Check with participants to be sure they aren't outside the sample (e.g. from higher grades)
- Properly train the supervisor and the assessors to select the students based on protocol.



# Monitor Interrater reliability (IRR) During Data Collection

- Conduct IRR checks with approximately 10% of the student sample
- Supervisor should
  - ensure assessors are following the IRR schedule and procedures
  - review disagreements that the assessors have identified
  - report disagreements up to the support team for comparison



# Working with Students

- *Make sure that waiting students cannot hear or see the testing. Ask the students to wait their turn in a space close to the testing area.*
- If the director prefers that the students wait in their classrooms, make sure those students do not leave school before they are tested.
- Write down names of sampled students before asking them to wait their turn. This will help identify the sampled students. Discard the list of names once assessment is completed for the school.
- **If a student does not agree to participate in EGRA, or is visibly distressed, the student is kindly dismissed and a replacement is selected.**

# Summary of Best Practices

- Know and follow all IRB guidance when working with children, and factor sufficient time into the work plan for IRB review and approval.
- Develop an implementation plan for full data collection in advance so that it can be fully tested during piloting.
- Ensure that all key players are engaged, know the purpose, and understand the intended output of the pilot.
- Plan for sufficient numbers of supervisors to allow one per school team. Select supervisors from the enumerator pool. Meet them immediately before data collection to verify all logistics and distribute materials.
- Plan based on the school calendar and any security, weather, or transportation concerns.
- Keep total data collection time to a minimum: 1–2 weeks is ideal.
- Obtain MOE approval and letters early and follow all government protocols without divulging sample schools/testing dates.
- If a student does not want to participate or is visibly distressed, she or he is immediately and kindly dismissed and an alternate selected.