



**USAID**  
FROM THE AMERICAN PEOPLE

# What will it take?

## Implementing assessments with impact

---

29 August 2024

Philippi Educational Measurement and Evaluation  
Association

2024 National Conference on Educational  
Measurement and Evaluation

# Overview of USAID Improving Learning Outcomes for the Philippines (ILO-Ph)

**Activity Goal:** This activity builds on long-standing support provided by the USAID to boost DepEd's capacity for conducting large-scale assessments, strengthening their data systems, and their strategic communications capabilities.

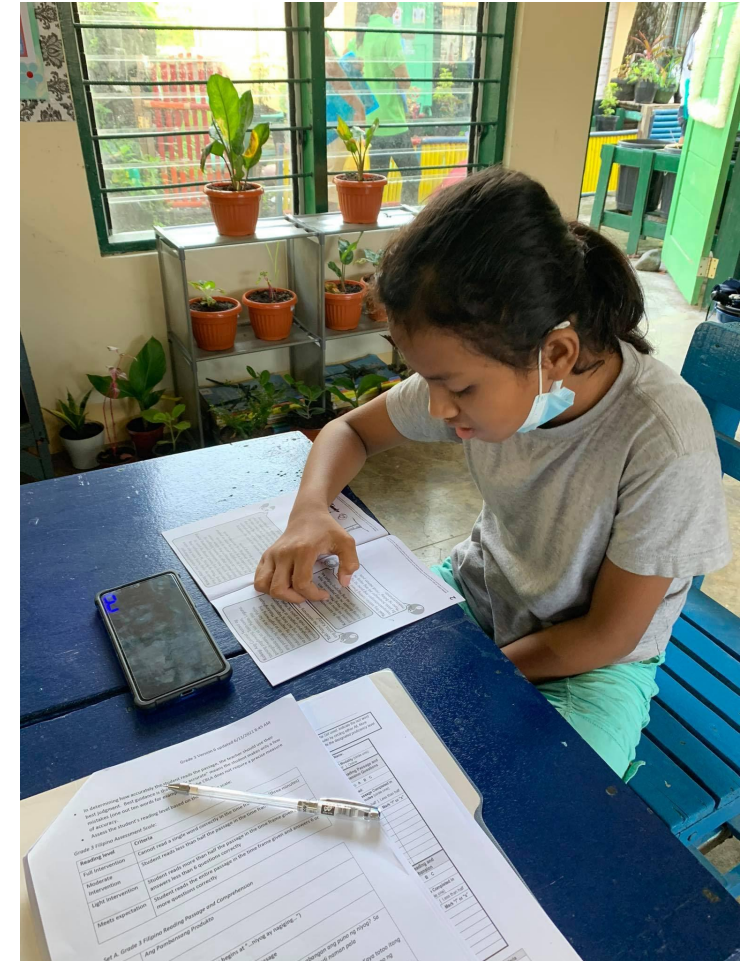
**Implementing Partners:** RTI International, SEAMEO-INNOTECH, Delivery Associates



# ILO-Ph Support to DepEd to Develop a Comprehensive Learner Assessment Framework

Partnered with DepEd/BEA and guided by the following questions:

- What is the inventory of all system and classroom-based assessments? How do they relate to or complement each other?
- How have the policy guidelines on system and classroom-based assessments evolved over time?
- What are the system capacity challenges for implementing assessments effectively and efficiently at national and subnational levels?
- What are the gaps and opportunities for strengthening inclusive assessment practices?
- What is the optimal assessment framework, given the various purposes, requirements, and system capacity challenges and opportunities?



# Methodologies



Policy documentation review;



Surveys of region and division testing coordinators;



Key informants interviews



FGDs with central, regional, division, and school-level educators;



Reflection of comparative international best practices.

## We sought to understand the challenges and opportunities in implementing assessments

+20

FGDs and interviews

*Leadership and project teams within DepEd, Education Administration experts, Assessment Design experts, Regional, Division and School-level staff*

5

Bureaus

*BEA, BLD, BCD, BLR, BAE*

3

regions

*Regions III, IV-A and V*

2

Stakeholder Forums

*DepEd External Stakeholder Forum and Regional and Division Stakeholder Forum consisting of Academe, Research Partners, Development Partners, and Region and Schools Division Leaders*

# We focused on the system's capacity...

## System capacity



### Assessment Design

The alignment and complementarity of different assessments



### Assessment Administration

The extent to which assessments are implemented effectively and efficiently



### Assessment Data

How data is used to inform learning interventions, policy, decision-making, and accountability.

**...spanning assessments at all levels.**

1

National Level System Assessments

2

Regional Level Assessments

3

Division Level Assessments

4

School Level Assessments

## There were five overarching takeaways

- 1 The **sheer number of assessments** at all levels **can be overwhelming**, affecting classroom instruction time and potentially leading to exam fatigue for students.
- 2 There is **duplication and limited coherence** among assessments, particularly at the subnational level.
- 3 The proliferation of assessments reflects **repeated attempts to address key data gaps**.
- 4 Even when data is available, it is **often not used effectively** to support student learning, nor drive accountability.
- 5 Moreover, the concern that **learners do not take system assessments seriously** undermine the credibility of the results.



# The sheer number of assessments at all levels can be overwhelming

Types of assessments	Examples of assessments	Degree of standardization
National-level system assessments	<ul style="list-style-type: none"> <li>•National Achievement Test (NAT)</li> <li>•National Learning Camp Assessment (NLCA)</li> <li>•Early Grade Reading Assessment (EGRA)</li> <li>•Early Language, Literacy, and Numeracy Assessment (ELLNA)</li> </ul>	Highly standardized; uniform for all students across the country.
Regional-level assessments	<ul style="list-style-type: none"> <li>•Regional Achievement Test (RAT)</li> <li>•Regional Diagnostic Assessment (RDA)</li> <li>•Regional Midyear Assessment (RMY)</li> <li>•Project All Numerates (PAN)</li> <li>•Science Process Skills Assessment (SPSA)</li> <li>•Comprehensive Rapid Literacy Assessment (CRLA)</li> <li>•Learning Assurance for Monitoring and Progress (LAMP) Assessment</li> </ul>	Highly standardized; uniform for all students across the region.
Division-level assessments	<ul style="list-style-type: none"> <li>•Periodic Quarterly Assessments</li> <li>•Summative Assessments</li> <li>•Formative Assessments</li> </ul>	Moderately standardized; Summative assessments are standardized within schools or districts, but degree of standardization may vary across the different divisions. Formative assessments are often tailored to the individual classrooms.
School-level assessments*	<ul style="list-style-type: none"> <li>•Formative assessments</li> <li>•Summative assessments</li> <li>•Quizzes</li> </ul>	No standardization; tailored to individual classrooms and teachers.

## What we heard (paraphrased)...

**Early career teacher**

*'I don't have any time to teach curriculum'*

**Principal**

*'The beginning of the year is busiest'*

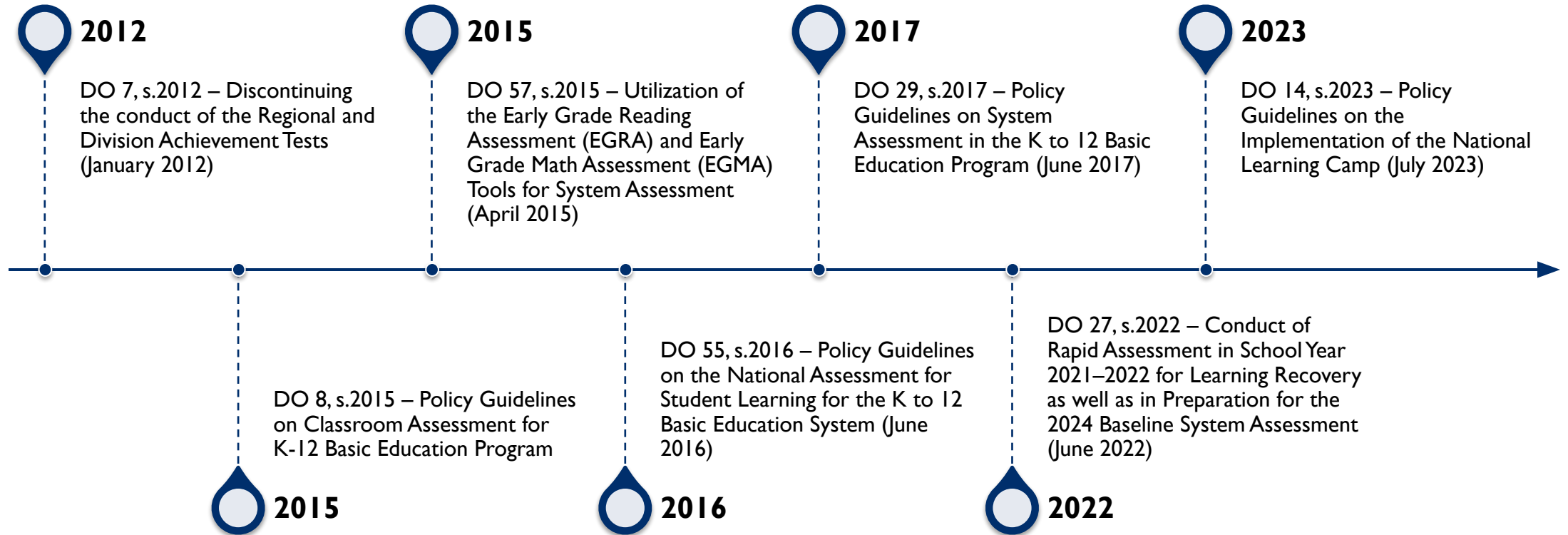
**Regional leader**

*'Students lose focus and aren't invested...they don't take it seriously'*

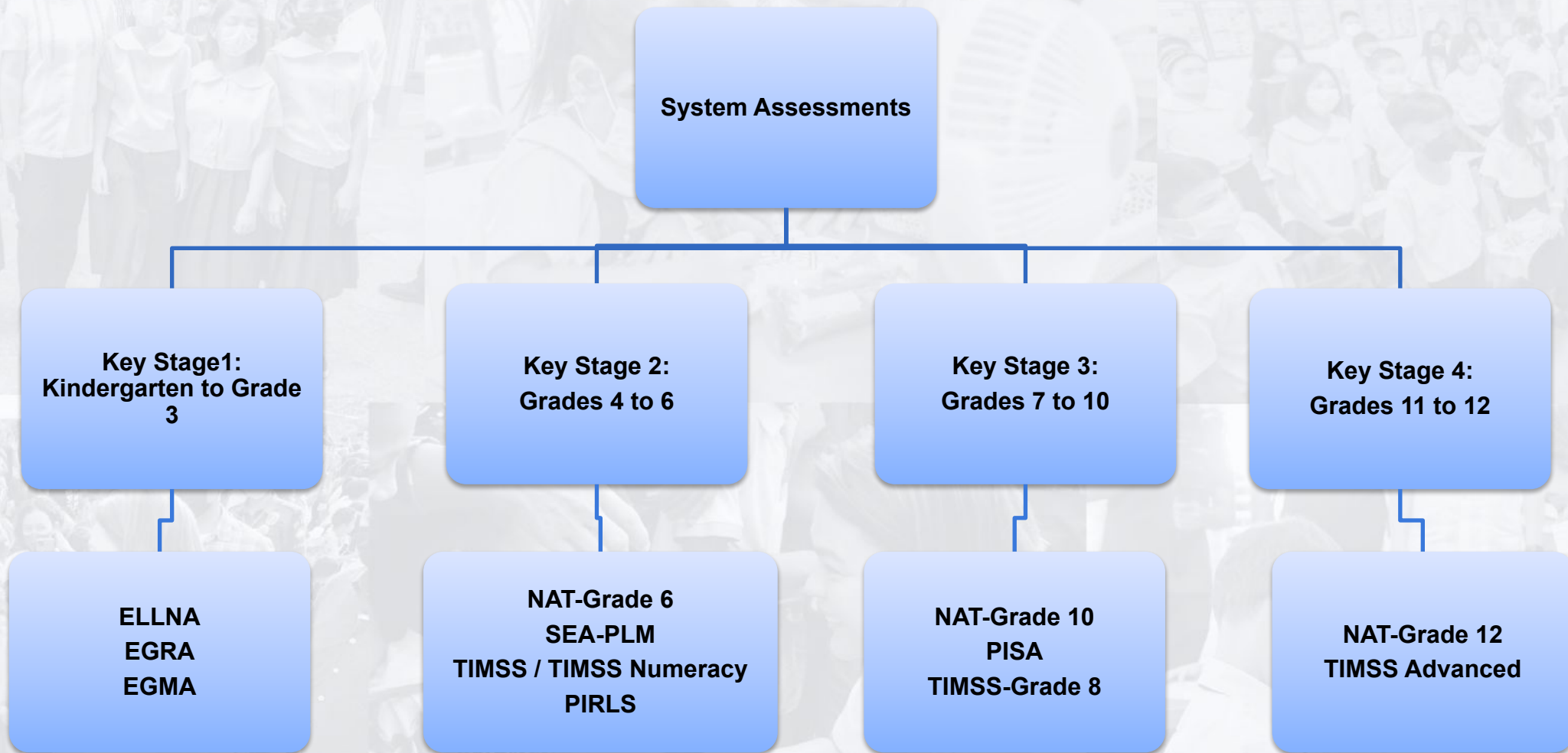
**Experienced teacher**

*'When am I meant to design, administer and follow-up all these assessments?'*

# Policy Framework Governing System and Classroom Assessments



# DepEd's System Assessments



# DepEd's 21<sup>st</sup> Century Skills Framework

**DepEd Order No. 55, s. 2016** (Policy Guidelines on the National Assessment of Student Learning for the K to 12 Basic Education Program).

**Holistically developed Filipino with 21<sup>st</sup> century skills.**



*Information, Media  
and Technology Skills*



*Learning and  
Innovation Skills*



*Effective  
Communication Skills*



*Life and  
Career Skills*

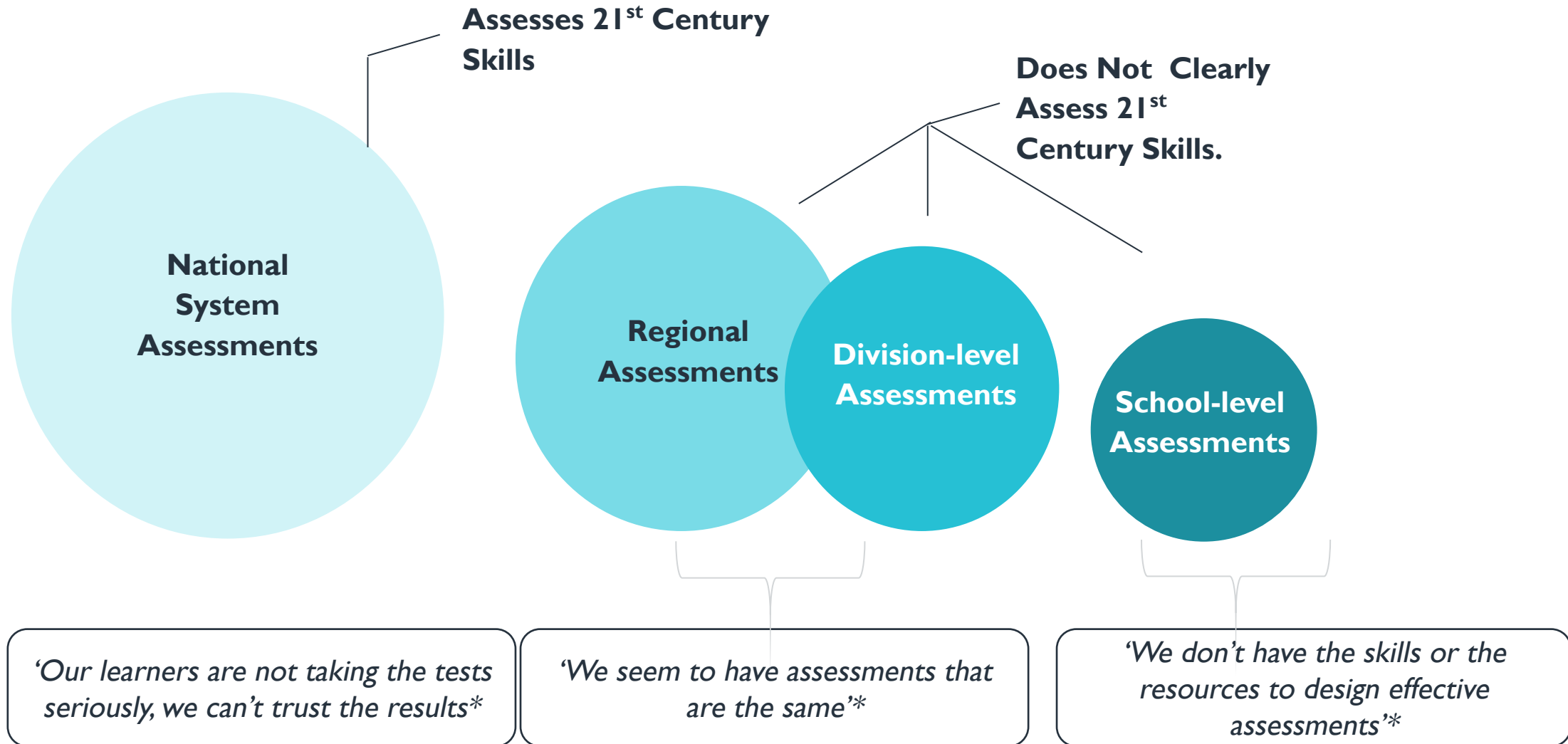
1. **21st-century Skills** are abilities embedded in the K to 12 Basic Education Curriculum that learners must acquire. These include:
  - a. **Communication Skills** refers to the ability to express one's self clearly and collaborate with others.
  - b. **Information, Media and Technology Skills** refers to the ability to gather, manage, evaluate, use, and synthesize information through media and technology.
  - c. **Learning and Innovation Skills** refers to the ability to think critically, analyze and solve problems, create and implement innovations, and generate functional knowledge.
  - d. **Life and Career Skills** refers to intrinsic and socialized personal values, ethics, and attitudes for life after basic education and learning within the workforce.

## Snapshot of Inventory of Region and Division-Initiated Assessments

Assessment Characteristics	Regional Result (%) (n=13)	Division Result % (n=111)
Respondents	14 out 17 respondents	96 out 145 respondents
Purpose	<b>Formative</b> = 2 out of 13 (15%) <b>Summative</b> = 6 out of 13 (46%) <b>Diagnostic</b> = 6 out of 13 (46%) <b>Graded</b> = 2 out of 13 (15%)	<b>Formative</b> = 39% (43) <b>Summative</b> = 49% (54) <b>Diagnostic</b> = 43% (48) <b>Graded</b> = 31% (34)
Includes all grade levels	<b>5 out of 13 (38%)</b>	All grade levels = 48% (53)
<i>Note: Most of the assessments also <b>focus on the grade levels at the end of each key stage</b> – Grades 3, 6, 10, and 12.</i>		
Covers all learning areas	7 out of 13 (54%);	Covers all learning areas = 50% (55)
<i>Note: Most of the assessments cover Mathematics</i>		
Covers 21 <sup>st</sup> Century Skills? Standardized?	Yes = 12 out of 13 (92%); No = 1 Yes = 12 (92%); Not sure = 1	Yes = 94%; No = 1%; Not sure = 5% Fully standardized = 65%; Quasi-standardized = 18% Not standardized = 13%
<i>Note: these figures are self-reported and not validated by BEA</i>		

Assessments of learning outcomes	Stage 1			Stage 2			Stage 3				Stage 4		Scope and frequency			
	K	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Sam ple	Cen sus	Frequency
<b>National measures of high-quality basic education</b>																
ELLNA				✓										✓	✓	Sampled yearly but census every 3 years
EGRA and EGMA				✓										✓		Sampled yearly or every 3 years
NAT Grades 6, 10, 12							✓				✓		✓	✓	✓	Sampled yearly but census every 3 years
<b>National-, regional- and schools division-initiated classroom-based assessments</b>																
CRLA/RMA or EGRA/EGMA		✓	✓	✓												Sampled twice yearly
Phil-IRI				✓	✓	✓	✓									Census, twice-yearly, English and Filipino
Region and Division Assessments		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Census, yearly, except for every 3 years when ELLNA/NAT are in census
District Unified Quarterly Assessments		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	Census, quarterly
<b>International benchmarks of high-quality basic education</b>																
PISA								✓	✓	✓	✓	✓		✓		Sampled 15-year-old learners, every 3 years
TIMSS					✓				✓					✓		Sampled every 4 years
SEA-PLM						✓								✓		Sampled every 5 years

# There is often duplication and limited coherence among assessments



# Data is the critical issue

Data is rarely shared...

Results from system and regional assessments are **often not shared** while data from Division and School level assessments are more likely to be readily available.

Additional assessments are rolled out to meet data needs

...and when it is shared, it's use is limited...

Results are often **shared late and do not provide the analysis** required to meaningfully deliver insights

Data is not being used to inform teaching practice or drive accountability (both to drive system-level and individual performance)

...representing a missed opportunity.

There are no silver bullets, **but if there was one...**

Educators are desperate for data to inform their teaching practice, and want to be benchmarked against their peers



## To whom are Region- and Division-initiated test results communicated

Results communicated to...	Region Results	SDO Results
Learners	0 out of 13 (0%)	49% (54)
Parents and caregivers	0 out of 13 (0%)	47% (52)
Teachers	4 out of 13 (31%)	83% (92)
School Heads	3 out of 13 (24%)	83% (92)
District and Division Supervisors	6 out of 13 (46%)	94% (104)
Regional Offices	3 out of 13 (24%)	20% (22)
Central Offices	9 out of 13 (69%)	1% (1)
Other stakeholders (e.g., LGUs)	1 out of 13 (8%)	10% (11)

Source: Inventory Survey of Region and Division Testing Coordinators: self reported

# Computer-based assessments could expedite access to data

## The opportunity

Computer-based assessments enable **quicker generation** of assessment results data than paper-based assessments.

## What we've observed

Some interviewees had yet to receive results from the latest computer-based NLCA, indicating other factors may be slowing down the sharing of assessment result data

## What we need to account for

The **necessary infrastructure** needs to be in place to ensure equitable access.

**Bureaucratic, administrative and technological factors were identified as key challenges in making data more readily accessible**

# Updating of Data Sharing and Utilization Guidelines

- The current regime limits access and use of system assessment results by external stakeholders
- The new Assessment policy will propose updates to the data sharing and utilization guidelines in the following ways:
  - Production of Assessment Result Dashboards for internal (to school level) and external access (to division level)
  - Removal of payment charges for external data requests
  - Removal of 10% data limit:
    - instead allow 100% of summarized school data nationally;
    - Allow 100% of anonymized learner-level data on a schools-division basis depending on purpose of research
  - Incorporation of an Ethics Review Committee
  - Adherence to a data processing timeline and timely response to data requests

# How do assessments fit within DepEd's Performance Management, Support and Accountability System?

We identified five key components that underpin a performance culture



**Outcome-focused targets and supporting indicators** that are well, aligned, understood and communicated



**System leadership** that prioritizes a culture of high-performance rather than compliance



**Quantitative data** that is **available and regularly shared** to understand system performance



**Agency** across the system to flexibly deploy plans and resources



**Targeted and tailored support** for system actors, including a focus on **support for underperformers**

With this view we identified a well-intentioned system struggling to support its workforce and its students



**Lack of focus:** 'if everything's a priority, nothing is a priority'



**Inadequate data:** insufficient to track progress and for data-informed support



**Compliance culture:** too many policies, processes and tools



**Misalignment:** limited alignment between policies and with the budget



**Lack of agency:** some regions feel disempowered

# Inadequate data: insufficient to track progress and for data-informed support

## Gaps

### 4. At a national level there is limited utility in the NAT...

Sample based so can't be used to hold different parts of the system to account, and limited data availability

### 5. ...and in regions, data is not consistently used to improve learning outcomes

Focus on output rather than outcomes (e.g. capital build, compliance related measures etc.).

**A system 'hoping for the best'**



## Recommended action

4. Implement 'fit for purpose' **standardized assessments** to track progress and to inform where support is needed most

5. Use data to **understand progress and drivers** of performance and align the targets to **focus technical assistance** to schools

6. Attach **higher stakes** to the outcome of system assessment results **for both learners and system actors** (organizationally, not individually)

**A system informed by facts**

# Assessment Gaps for Enhanced Performance Management, Support and Accountability System (EPMAS)

**System assessments** are not fit for purpose:

- Measurement – focus on 21<sup>st</sup> Century Skills vs Learning Domains
- Timeliness and granularity of results – Too late to direct technical support to schools
- Not consistently census-based (for ELLNA, G6 and G10) or consistently implemented
- Students do not take the assessments seriously
- Policy guidelines do not explicitly link assessments to results-based management system policies (e.g. performance-based bonuses, organizational performance and commitment forms)

**Region and division-initiated assessments** are flawed

- Inconsistent administration across regions and schools divisions
- Resource and time-intensive
- Teacher-administered, introduces possible bias
- Results not comparable across regions and/or divisions
- Timeliness and utilization of results are limited

# System Assessment Requirements for Effective EPMAS

How to put in guardrails against unintended consequences of potential higher stakes:

- **Higher stakes** so that learners take it seriously- a true exit exam and/or feeder to remediation
- **Emphasize organizational accountability** vs individual accountability for assessment results
- **Content** must reflect the minimum and most essential learning competencies (MATATAG)
- **Timing:** assessments should take place in 2<sup>nd</sup> Quarter to get results in time
- **Balance indicators** so that teachers don't teach to the test or schools don't game the system
- **Census-based**, or at the very least, every school participates
- **Computer-based** (to ease administration and facilitate data processing)
- **Decentralized** data processing and data analysis (capacity building)
- **External or third-party** (private) overseen by BEA to ensure integrity of administration
- **Classroom-based** formative and summative assessment practices must be mainstreamed into teachers' normal instructional practices, but not used as tools to monitor system performance



# For discussion: your reflection...

- 1 How can we **incentivize learners (and schools) to take assessments more seriously**, particularly National-system level assessments?
- 2 Can we **explore computer-based options to make assessment results data more readily available** and meaningful?
- 3 Can we **strengthen the capacity of regions, schools divisions, and teachers to develop assessments** by recruiting relevant experts and provide training on assessment development to its existing staff? For example, incorporating 21<sup>st</sup> Century and higher order skills in assessments.
- 4 What types of **technical assistance or support could regions** and schools divisions benefit from to strengthen their design, administration and use of assessment data? For example:
  - testing the test items (eg, psychometric analysis to determine reliability and validity of measures)
  - item banks that are drawn from items developed locally as well as nationally
  - data systems to enable access and timely analysis of results?
  - exploring the role of BEA in supporting regions and divisions in these endeavours

THANK YOU!



**USAID**  
FROM THE AMERICAN PEOPLE