Educational Improvement, the Role of Assessment and Capacity Building

Dr John Cresswell
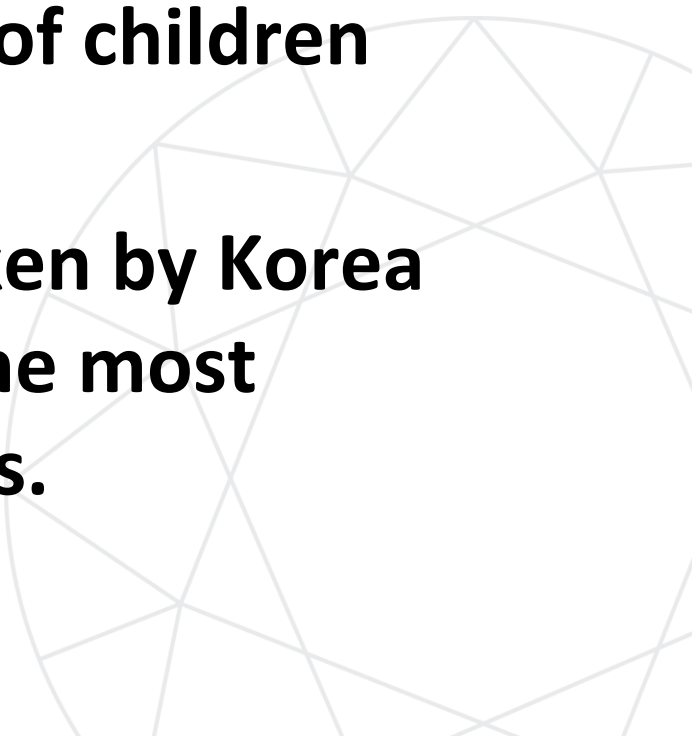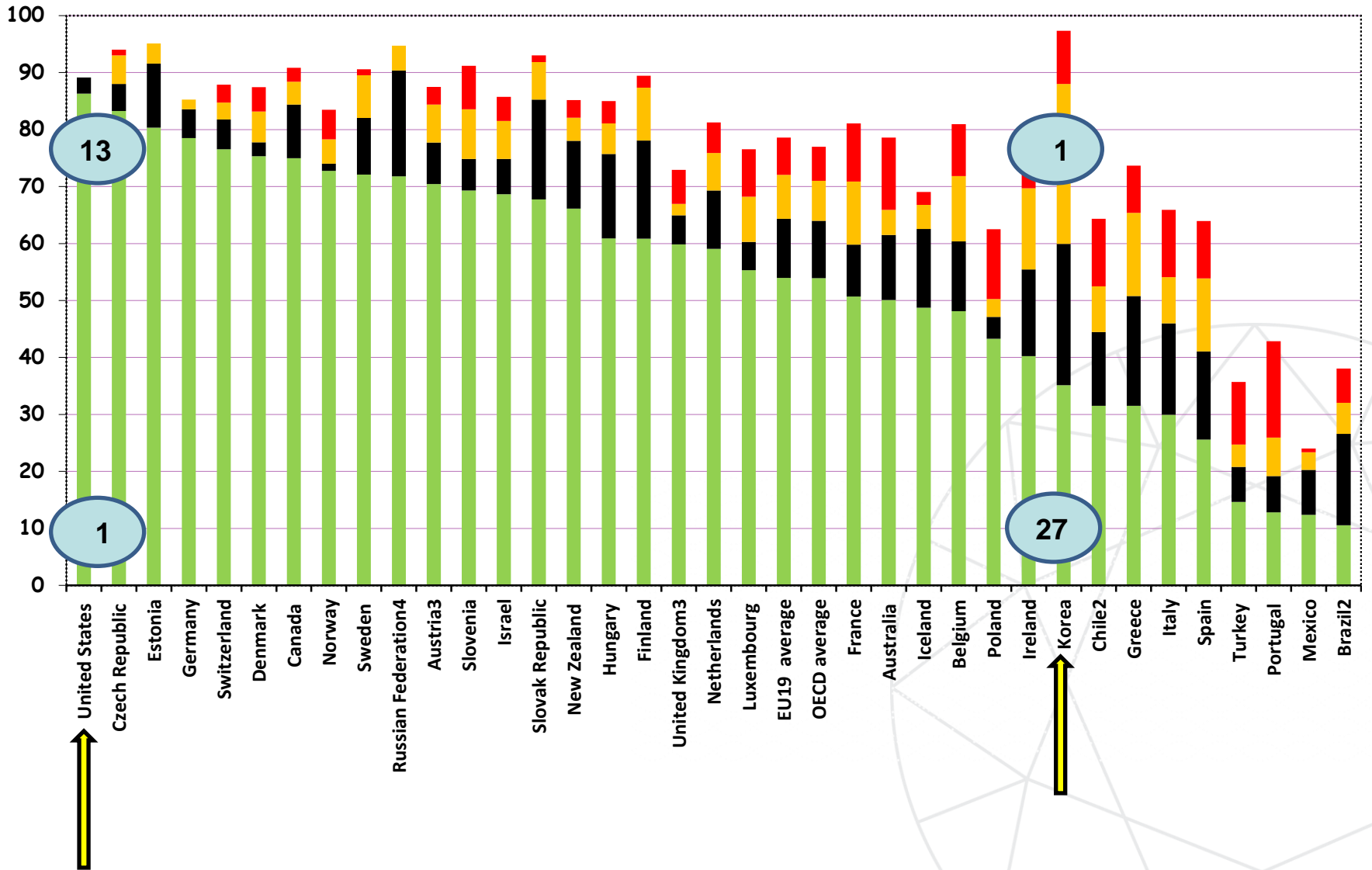
Siem Reap

August 21, 2014

Change happens

- **In Korea, in the 1950s, the literacy rate was 22% and less than 20% of children attended secondary school.**

- **Education reforms undertaken by Korea led to them having one of the most advanced education systems.**
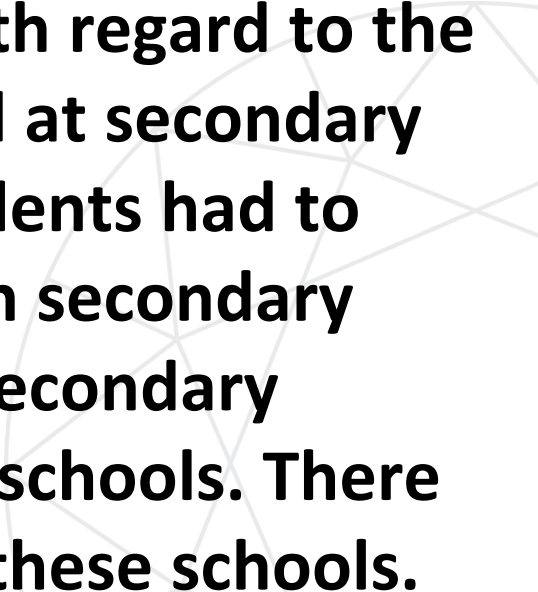
Approximate percentage of adults with high school qualifications

■ 1990s   ■ 1980s   ■ 1970s   ■ 1960s
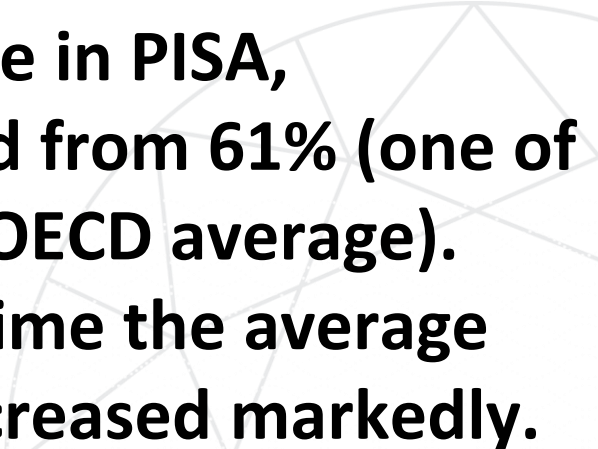
Change happens - Poland

- **In 1999, the Polish education authorities decided to make major changes to the structure of the education system, especially with regard to the number of years students spend at secondary school. Prior to the change, students had to follow three different streams in secondary schooling - general secondary, secondary vocational and basic vocational schools. There were vast differences between these schools.**
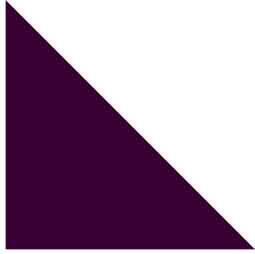
Change happens - Poland

BEFORE THE REFORM OF 1999

Grade				Age
0	Zero Class (Primary Schools or Kindergartens)			6
I	Comprehensive Primary Schools			7
II				8
III				9
IV				10
V				11
VI				12
VII				13
VIII	Exam	Exam		14
I	General Secondary Schools	Secondary Vocational Schools	Basic Vocational Schools	15
II				16
III				17
IV				18
V				19

Jakubowski, M., H.A. Patrinos, E.E. Porta and J. Wisniewski. 2010. "The impact of the 1999 education reform in Poland." World Bank Policy Research Working Paper No. 5263, Washington DC
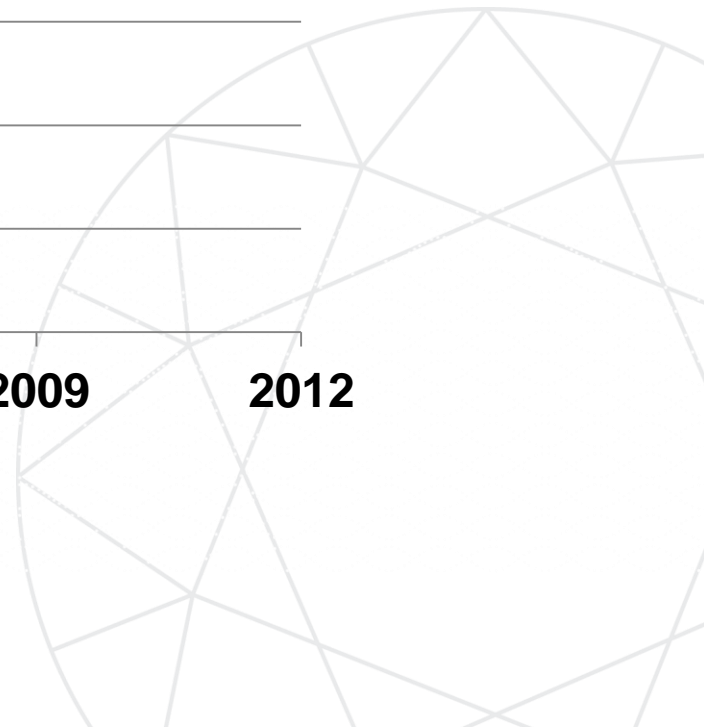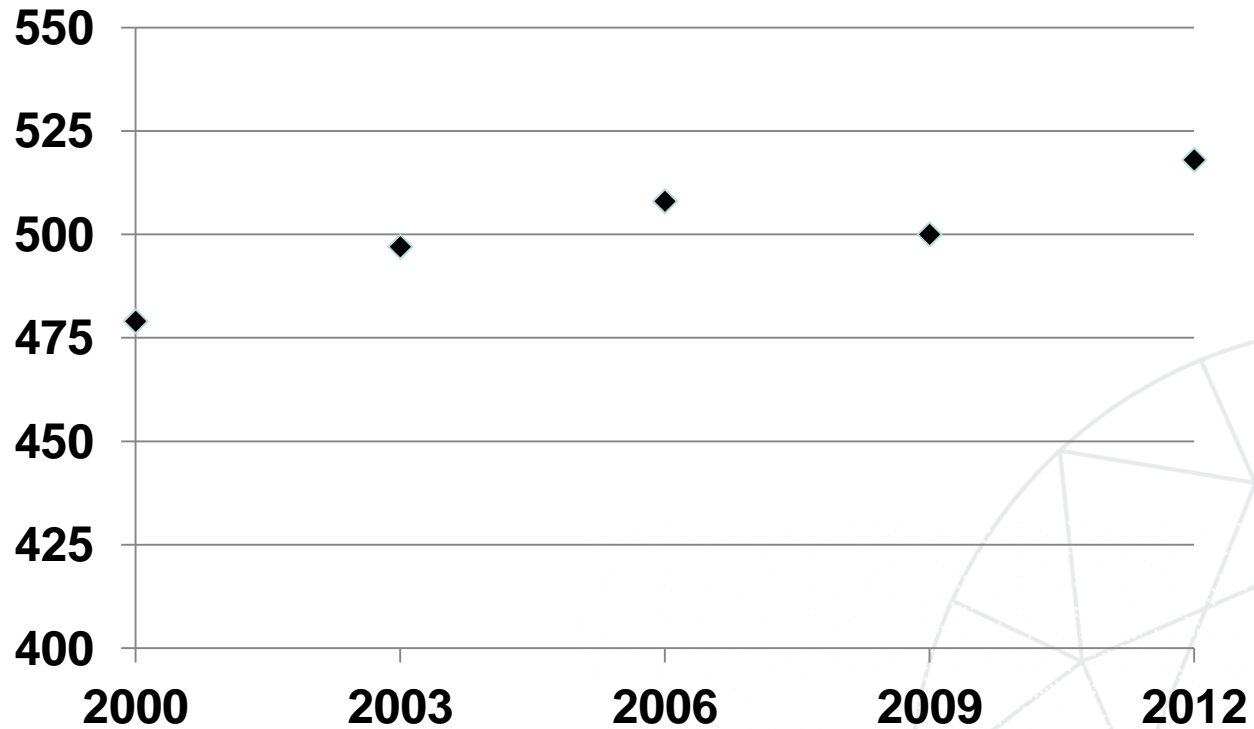
Change happens - Poland

- **This meant at age 15 (the PISA age) students were now in the same type of school. The amount of variation in student performance in PISA, explained by the school dropped from 61% (one of the highest) to 20% (below the OECD average). More importantly at the same time the average performance of the students increased markedly.**
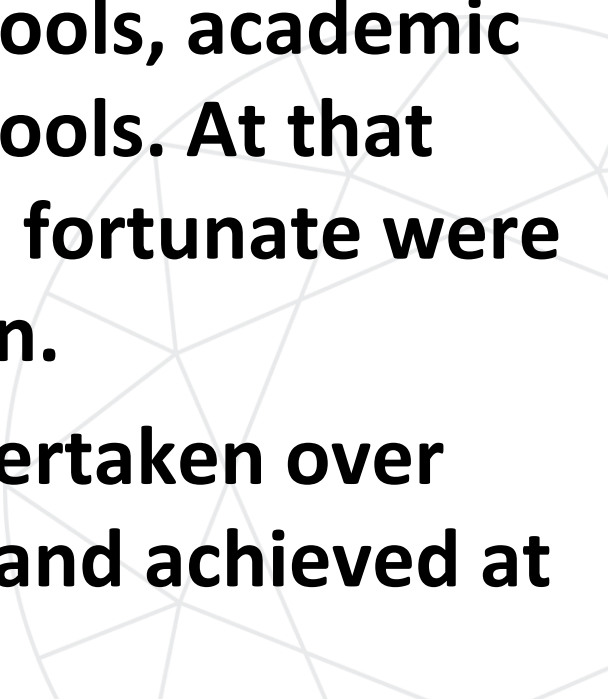
PISA results - Poland

Reading Score

Change happens - Finland

- **Until the late 1960s, in Finland most children left public school after six years. Some children went to private schools, academic grammar schools or folk schools. At that time only the privileged and fortunate were exposed to quality education.**

- **A series of reforms was undertaken over many years and in 2000 Finland achieved at the highest level in PISA.**

Reform cycle
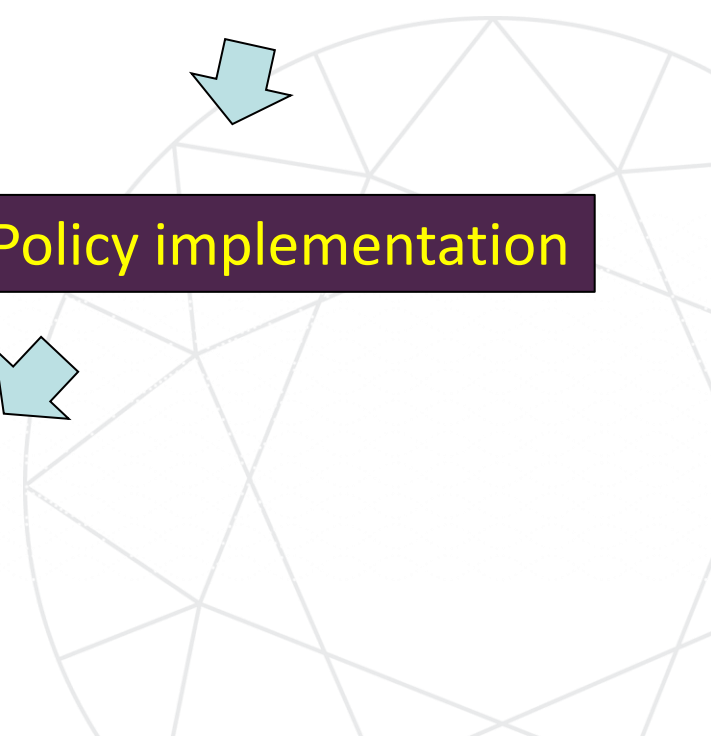
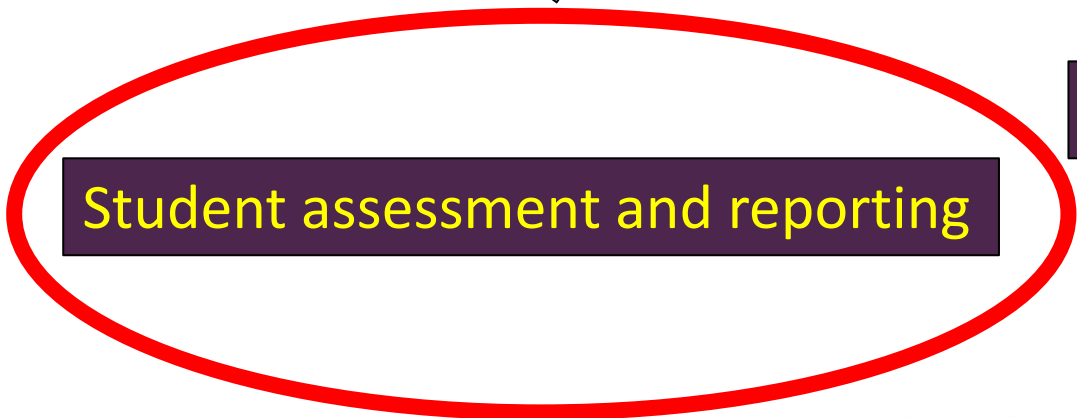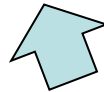**Education reform is not a straight line activity**

Reform discussion

Policy dialogue and decisions

Student assessment and reporting
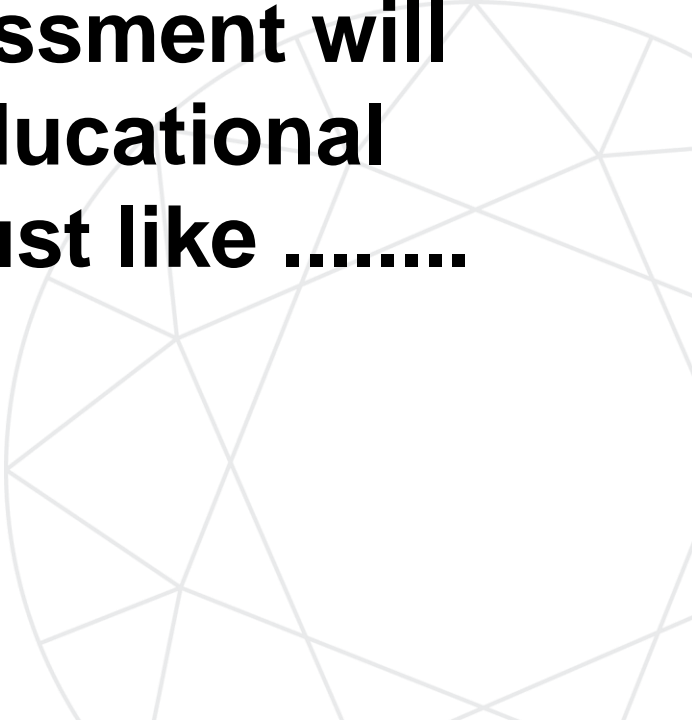
Policy implementation

Teaching and learning
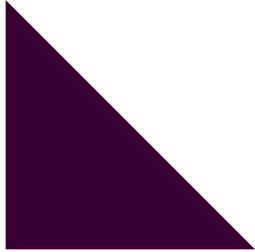
Reform cycle

- **Assessment is one part of the process of educational improvement.**
- **Simply doing more assessment will not, by itself, improve educational outcomes of students, just like ....**

.... you don't make a pig fatter by just weighing it

To make sure that an assessment program contributes to policy reform and improved student outcomes, it should have a number of elements
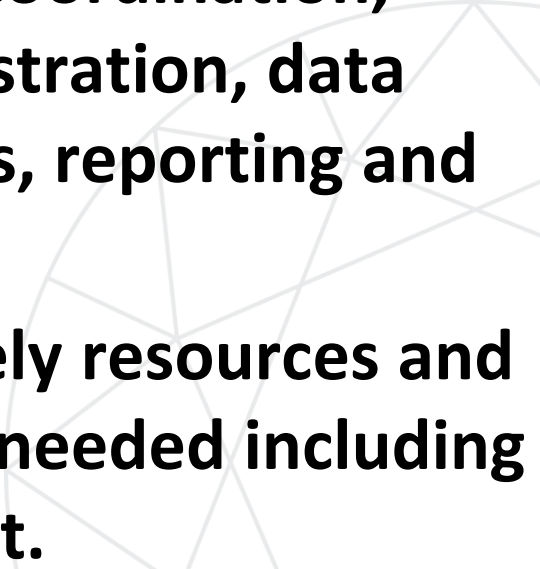
Elements of a robust assessment program
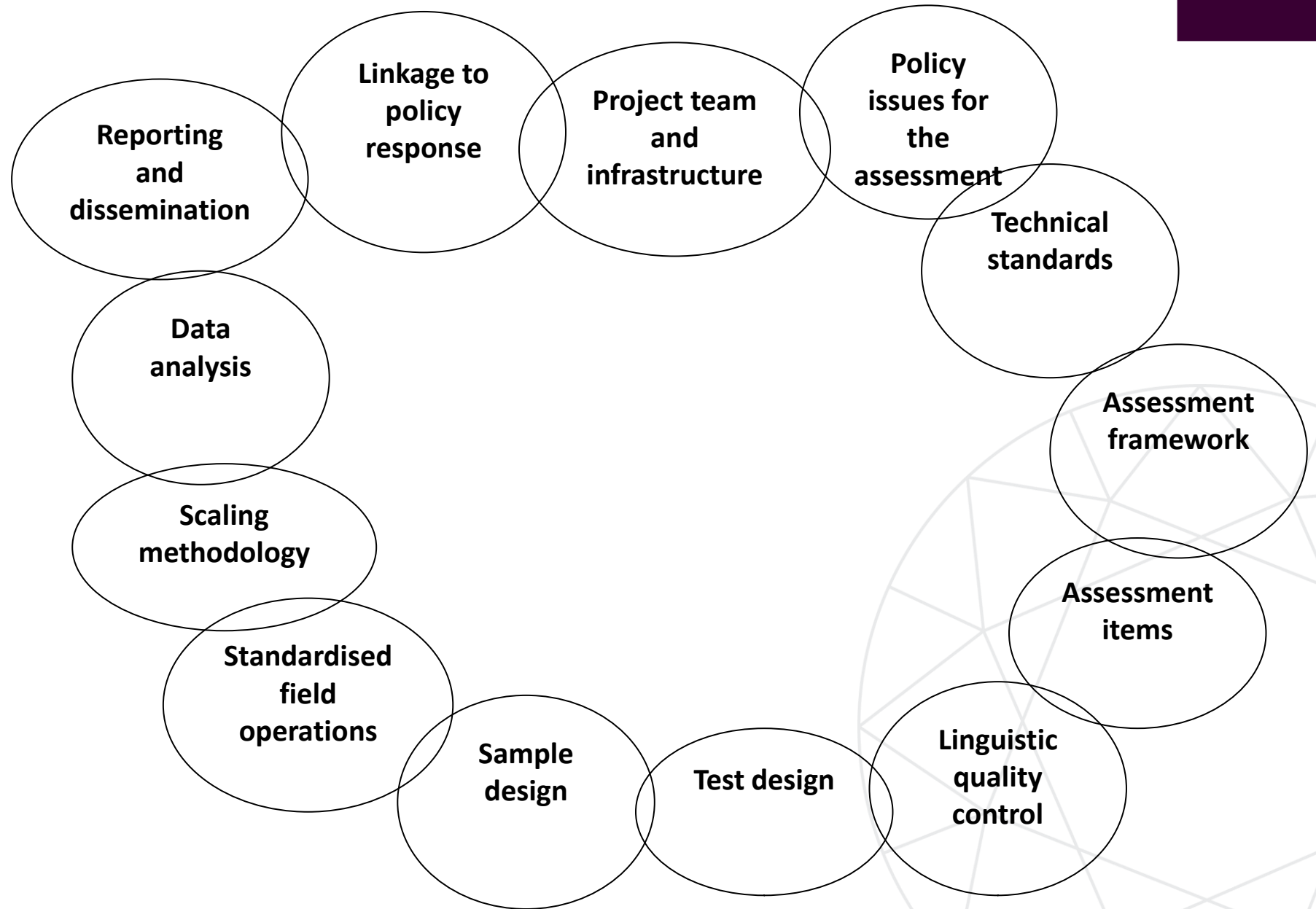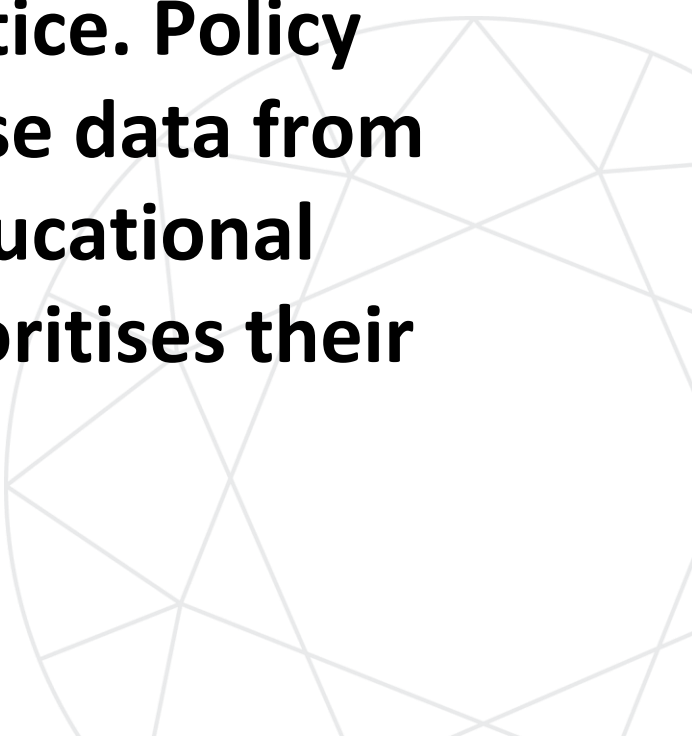
Project team and infrastructure

- **Staff are required to cover the roles of project management, policy liaison, assessment and item development, translation coordination, field operations, project administration, data management, sampling, analysis, reporting and communications.**

- **In addition to staffing, adequately resources and a core physical infrastructure is needed including computing and software support.**
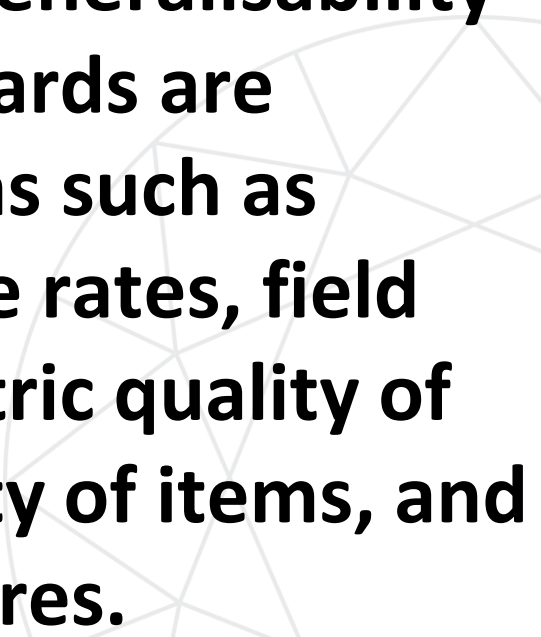
Policy issues for the assessment

- **The goal of an assessment is to estimate student outcomes in a way that informs educational policy and practice. Policy makers are more likely to use data from an assessment to inform educational policy if the assessment prioritises their policy concerns.**
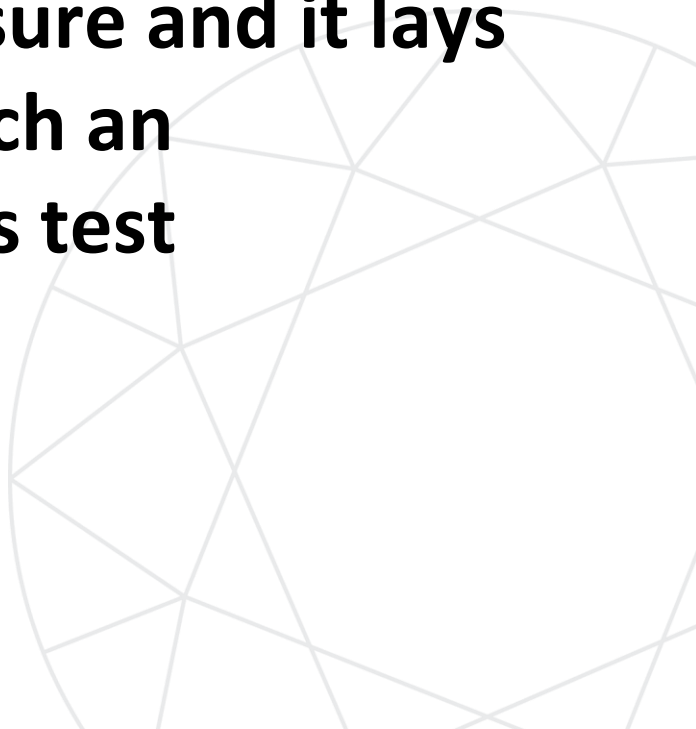
Technical standards

- **Technical standards reinforce the consistency, precision and generalisability of the data. Typically, standards are created for operational areas such as school and student response rates, field trial sample size, psychometric quality of the instrumentation, security of items, and test administration procedures.**

Assessment framework

- **An assessment framework is an explicit statement and discussion about what an assessment intends to measure and it lays out the principles upon which an assessment is built. It guides test development.**

High quality assessment items

- **Assessment items are the heart of an assessment program and careful attention must be paid to their development and quality. Item development is both a science and an art.**

- **If possible, field trial items with the target population.**
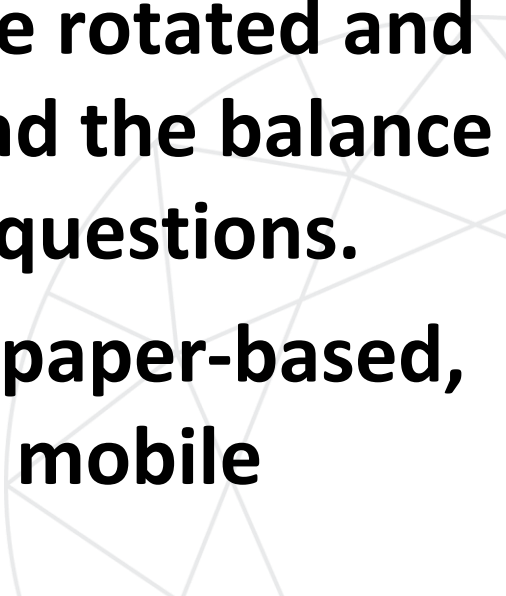
Linguistic quality control

- **Where assessments are administered in different languages, the different versions should be equivalent.**

- **Double independent translation is recommended.**
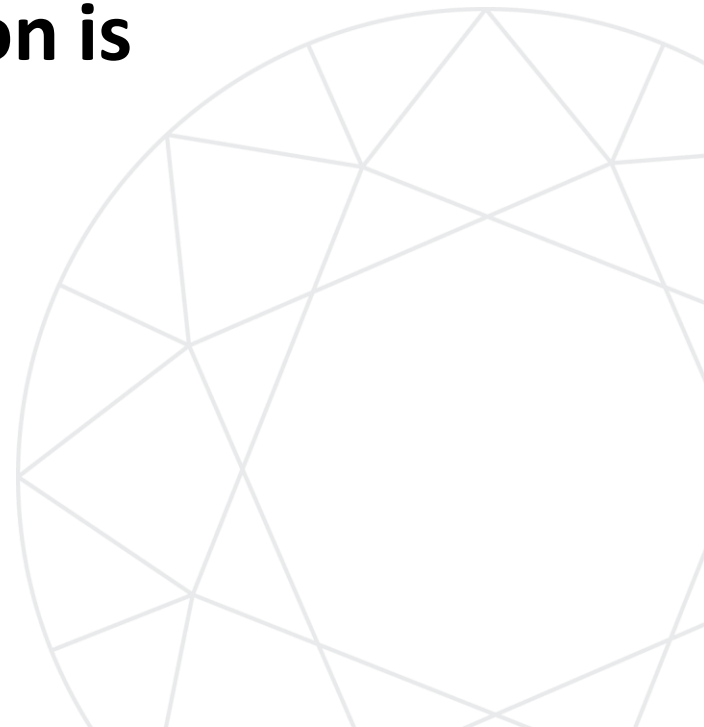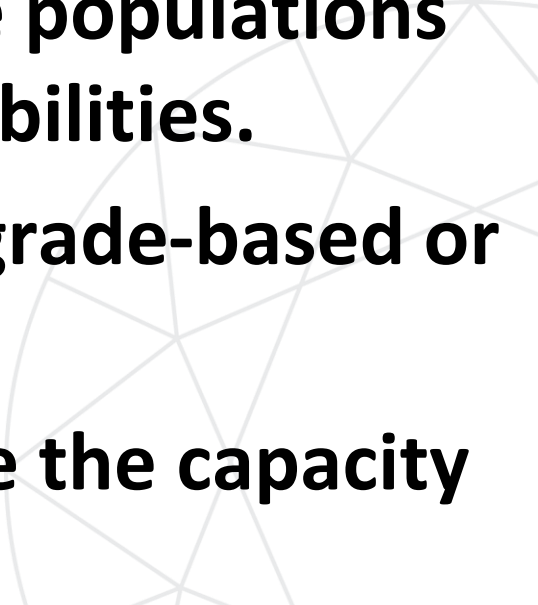
Test design

- **The test design will usually specify the format of materials; the number of test forms; how the materials are rotated and linked between booklets; and the balance of contextual and cognitive questions.**

- **Is the test computer-based, paper-based, or delivered on a tablet or a mobile phone.**

Sample design

- **This requires careful specification of target population(s) and techniques for accessing random samples from those populations with known sampling probabilities.**

- **Is it sample or census? Is it grade-based or age-based?**

- **Does your organisation have the capacity to do the weighting?**

Standardised field operations

- **Typically, field operations manuals are produced which specify the length of the testing sessions, the instructions for the students, and the physical aspects of the testing environment.**

- **Ideally, the test administrators will be independent and therefore impartial with regards to student outcomes.**

Scaling methodology

- **Item-response theory is currently recognised as the preferred methodology for evaluating and ensuring that the assessment instruments are of appropriate technical quality.**

- **Scales are constructed from the assessment data using item-response theory.**

Item Map

More capable students

More difficult items
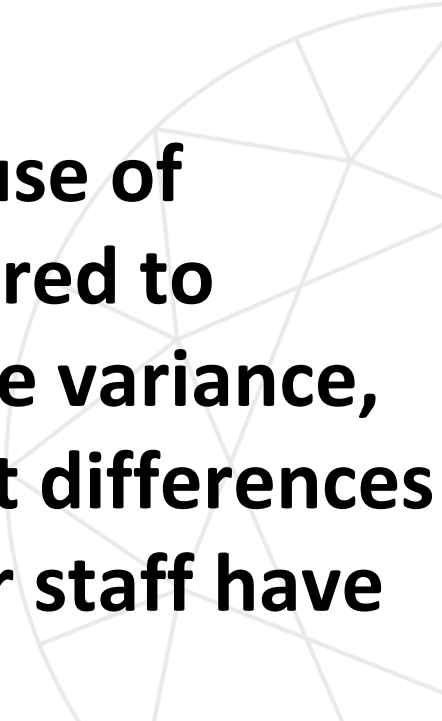
Less capable students

Easier items

Each X represents 103 students

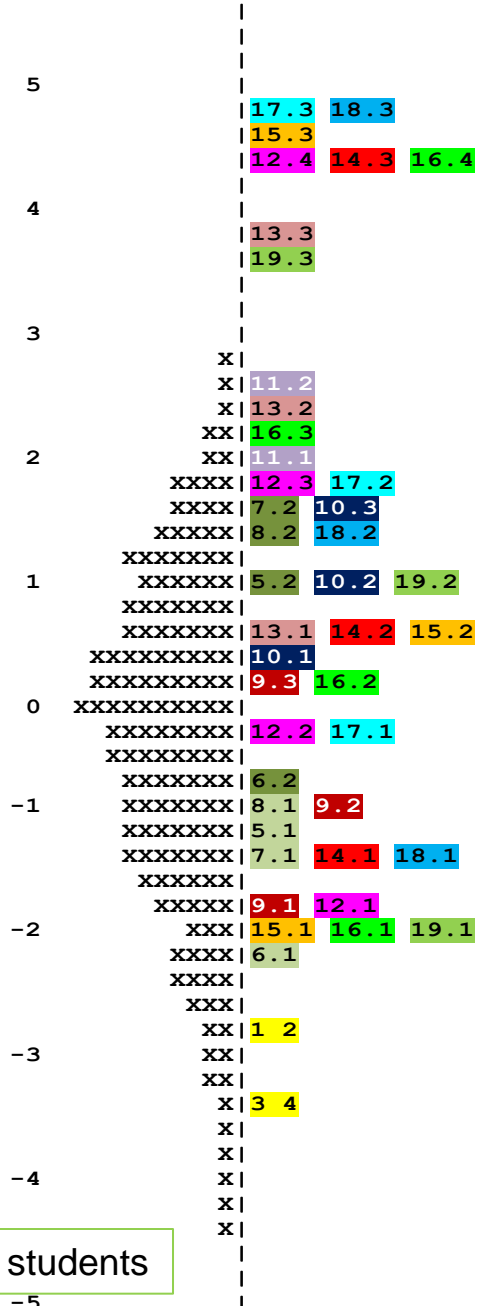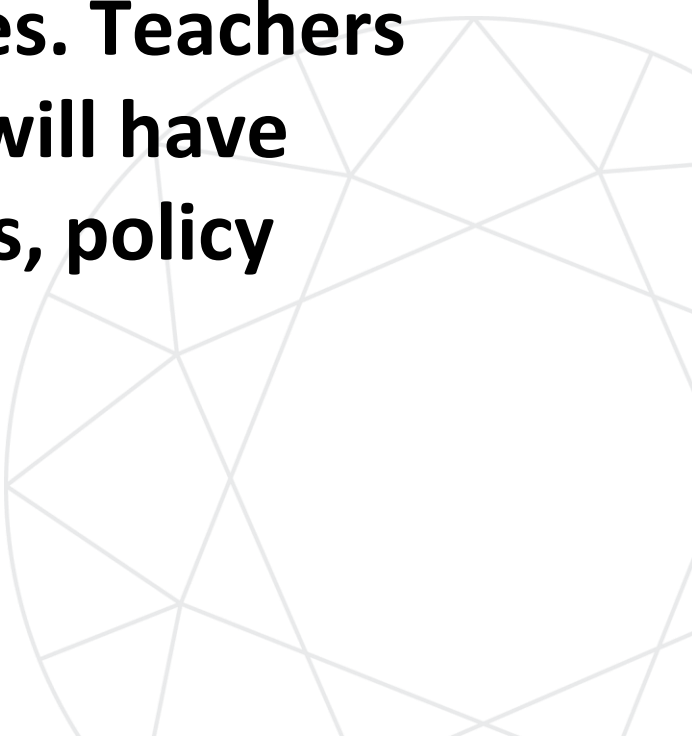|		
5	17.3 18.3	
	15.3	
	12.4 14.3 16.4	
4		
	13.3	
	19.3	
3		
X		
X	11.2	
X	13.2	
XX	16.3	
2 XX	11.1	
XXXX	12.3 17.2	
XXXX	7.2 10.3	
XXXXX	8.2 18.2	
XXXXXXX		
1 XXXXXX	5.2 10.2 19.2	
XXXXXXX		
XXXXXXX	13.1 14.2 15.2	
XXXXXXXXX	10.1	
XXXXXXXXX	9.3 16.2	
0 XXXXXXXXXX		
XXXXXXXX	12.2 17.1	
XXXXXXXX		
XXXXXXX	6.2	
-1 XXXXXXX	8.1 9.2	
XXXXXXX	5.1	
XXXXXXX	7.1 14.1 18.1	
XXXXXX		
XXXXX	9.1 12.1	
-2 XXX	15.1 16.1 19.1	
XXXX	6.1	
XXXX		
XXX		
XX	1 2	
-3 XX		
XX		
X	3 4	
X		
X		
-4 X		
X		
X		
-5		

Data analysis

- **Using item-response theory for scaling, the calculation of plausible values is the standard for achieving unbiased population estimates.**

- **With complex surveys, the use of replication methods is required to accurately determine sample variance, and consequently significant differences between groups – does your staff have this capacity?**
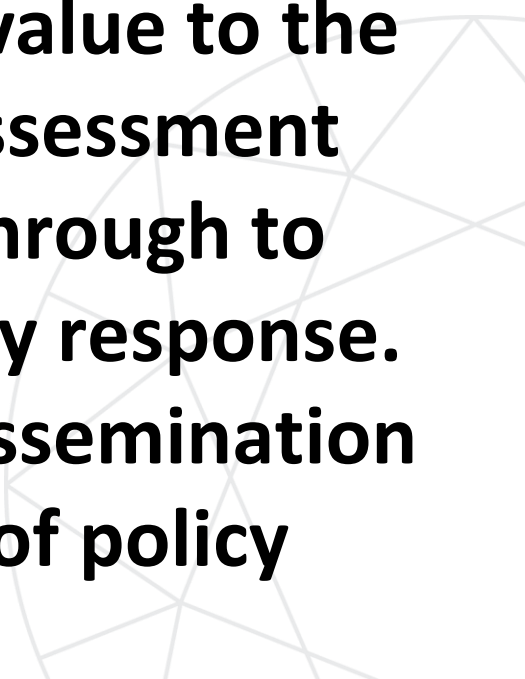
Reporting and dissemination

- **It is important to determine an appropriate strategy for disseminating the results to different audiences. Teachers and curriculum developers will have different foci, as will parents, policy analysts and economists.**

Linkage to policy response

- **Relating the assessment back to the policy issues determined at the outset of the development process adds value to the program. A well designed assessment program includes a follow through to inform and influence a policy response. This activity goes beyond dissemination and can include an analysis of policy options.**

In summary

If learning assessments are implemented, they should involve the systematic and strategic collection of data on education outcomes and factors related to these outcomes and be used to inform policy and practice aimed at improving education for all learners

We need data on outcomes – in addition to inputs

Focused on where learners are at, not where we want them to be

regularity & sustainability

In summary

If learning assessments are implemented they should be **USE of the data** matic and strategic collection of data on education outcomes and factors related to those outcomes and be used to inform policy and practice aimed at improving educational progress for all learners

and growth

Instrumentation and methodology better oriented to informing practice and policy

In summary

- **Improvement does occur in education, but it requires planning, resourcing and patience.**

- **Assessing student outcomes is an important step in the Reform Cycle.**

- **There are 13 elements that contribute to a robust assessment program.**

- **Ideally countries can build the capacity of their staff to achieve these elements.**

In summary

In the post 2015 education debate the paradigm is shifting from "access to education" to "access *and* learning".

There will be a greater focus on assessment of student learning outcomes.

On the first day of this meeting it was said that this was like saying …
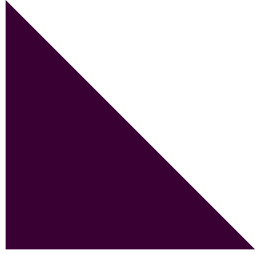
The proof of the pudding is in the eating ….

... or as they say in Cambodia

"The proof of the amok is in the eating"

- delicious
- ឆ្ងាញ់
- 美味
- स्वादिष्ट
- lezat
- स्वादिष्ट
- ਸੁਆਦੀ

- ருசியான
- lazat
- délicieux
- لذيذ
- مزيدار
- อร่อย
- ngon

Very sorry if I missed your language

Thank you

john.cresswell@acer.edu.au