United Nations
Educational, Scientific and
Cultural Organization

Regional Capacity building workshop on alignment between curriculum, teacher training and learning assessments

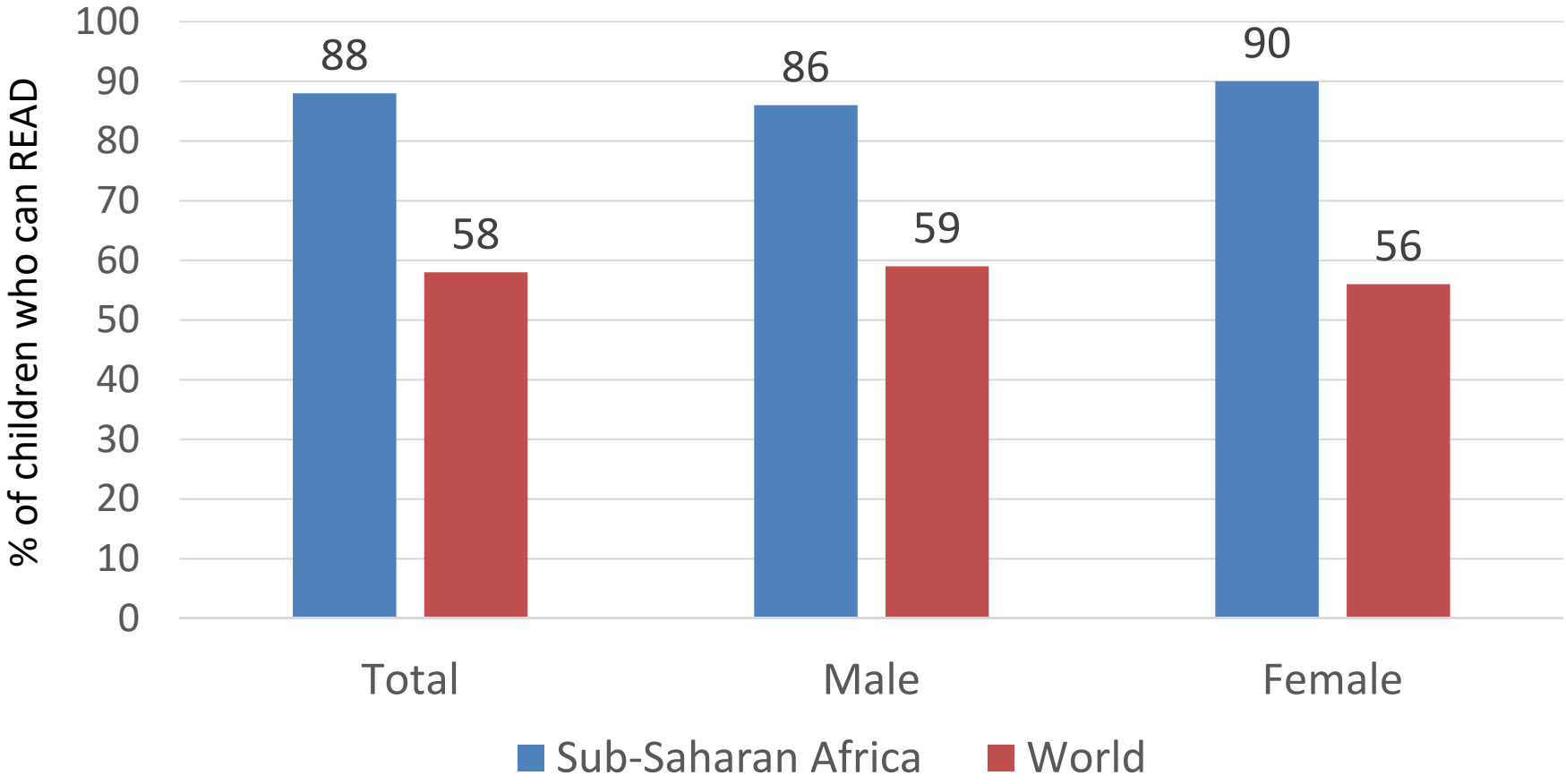
Dar es Salaam

25 July 2018

Aligning curriculum, teacher-training and learning assessment: the role of large-scale learning assessments

Education 2030

More than half of children and adolescents are **NOT LEARNING THE BASICS**



Source: UNESCO Institute for Statistics. 2017. More than one half of children and adolescents are not learning worldwide. Fact Sheet No. 46

Sustainable Development Goal 4
Objectif de Développent Durable 4

Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

Assurer une éducation inclusive et équitable de qualité et promouvoir des possibilités d'apprenstissage tout au long de la vie pour tous

SDG4 Outcome Targets:
Effective and Relevant Learning Outcomes

- ✓ Relevant and effective learning outcomes in primary and secondary education [4.1]

- ✓ Readiness for primary education [4.2]

- ✓ Technical and vocational skills for employment, decent jobs & entrepreneurship [4.4]

- ✓ Youth and adult literacy and numeracy [4.6]

- ✓ Competencies for global citizenship and sustainable development [4.7]

No shortage of large-scale learning assessment in Sub-Saharan Africa

- Half (28/54 or 52%) have their own **national learning assessment**
- 41% have participated in the latest cycle of **a cross-national assessment**

… in addition to public examinations AND classroom/formative assessments for informing teaching strategies

However, these **DO NOT** contribute to monitoring learning achievement at the global level

Source: Compiled from UIS Database of Learning assessments, 2011 ISCED Classifications, GMR 2015 and EPDC mapping (FHI 360, 2015)

Cross-national learning assessments

- Conducted in more than one country
- Measure cross-curricular knowledge, skills and competencies in a limited number of domains
- Provide feedback on the comparative performance of the system at particular grade or for a particular age or age-group
- Generally sample-based
- Uniform and standardized in terms of content, administration process, timing, scoring and analysis

Cross-national learning assessments around the world

- ✓ PISA Programme for International Student Assessment
- ✓ TIMSS Trends in Mathematics and Science Survey
- ✓ PIRLS Progress in International Reading Literacy Study
- ✓ ICCS International Civics & Citizenship Education Study
- ✓ ICILS International Computer and Information Literacy Study

Regional assessments

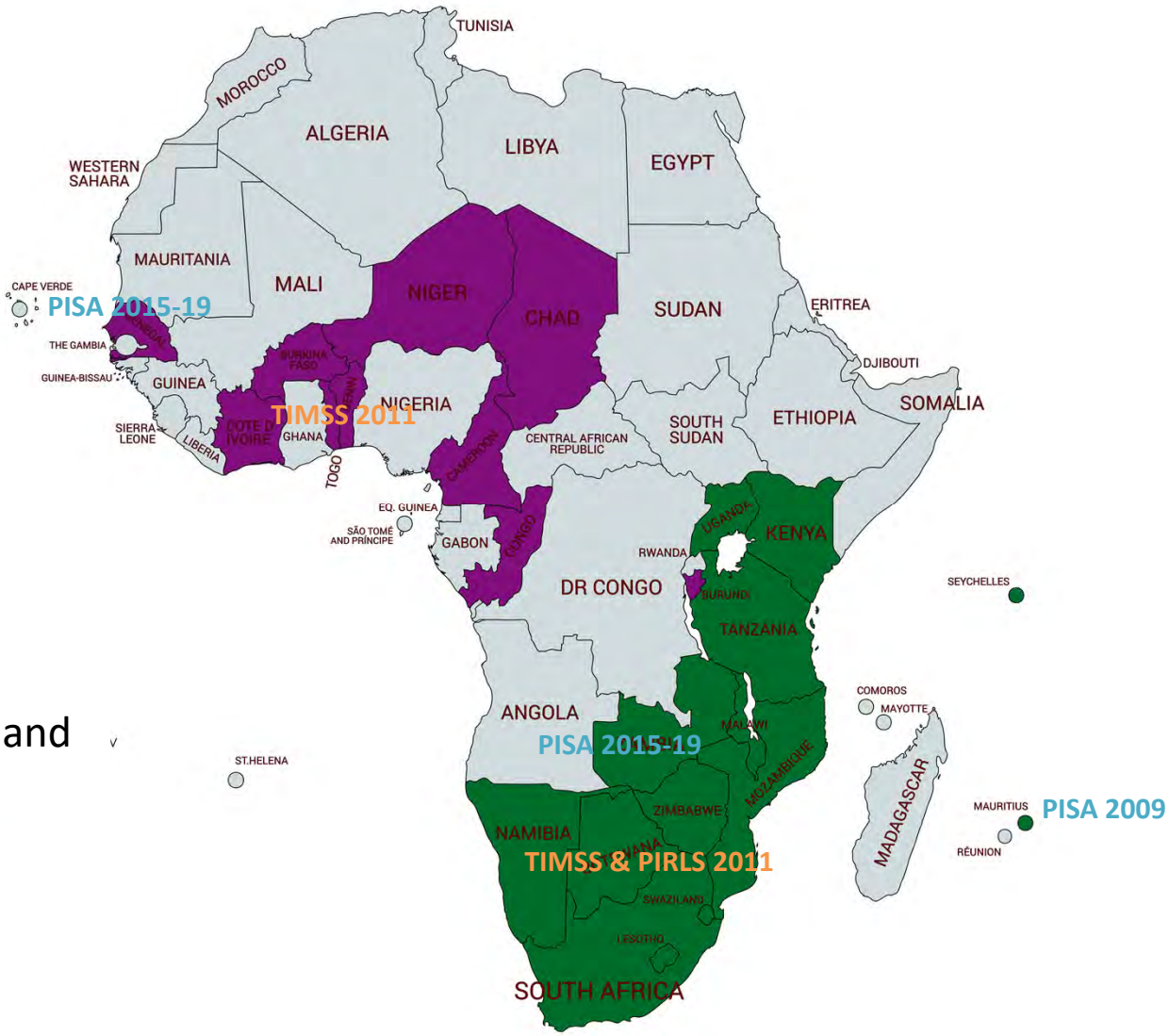
- ✓ PASEC and SEACMEQ Sub-Saharan Africa
- ✓ LLECE Latin America
- ✓ SEA-PLM and PILNA South East Asia / Pacific Islands

REFLECTION

Why do you think that cross-national assessments are important to consider for the alignment of teacher training, curricula and learning assessment?

Pourquoi pensez-vous qu'il est important de prendre en compte les évaluations transnationales pour l'alignement de la formation des enseignants, des curricula et de l'évaluation des apprentissages?

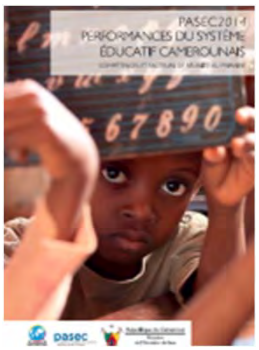
Cross-national learning assessments in Sub-Saharan Africa

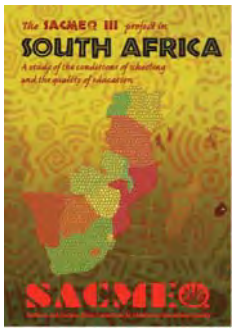


- SEACMEQ III and
- PASEC 2014

Cross-national learning assessments: the international report and national reports

- **Describe the knowledge** and skills of a target population;
- **Highlight disparities in learners'** (sub-populations') cognitive **abilities** by important socio-economic, regional, gender or other dimensions, including migration status and mother tongue;
- **Understand the factors that influence learning** achievement [e.g. home and school context and practices], and if these are changing over time; and
- **Identify general trends** in learning achievements and evaluating progress towards specific targets, using a set of indicators.





Girls perform significantly better than boys, particularly at the higher grades in Burundi

Figure 3.4: *Mathematics Performance Gap between Girls and Boys – Early Primary*

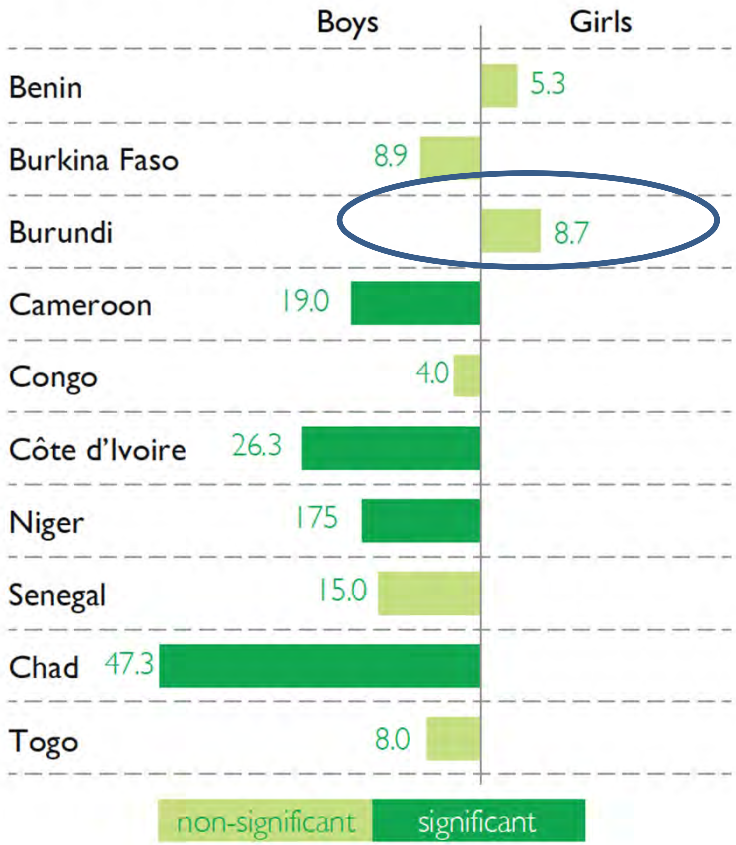
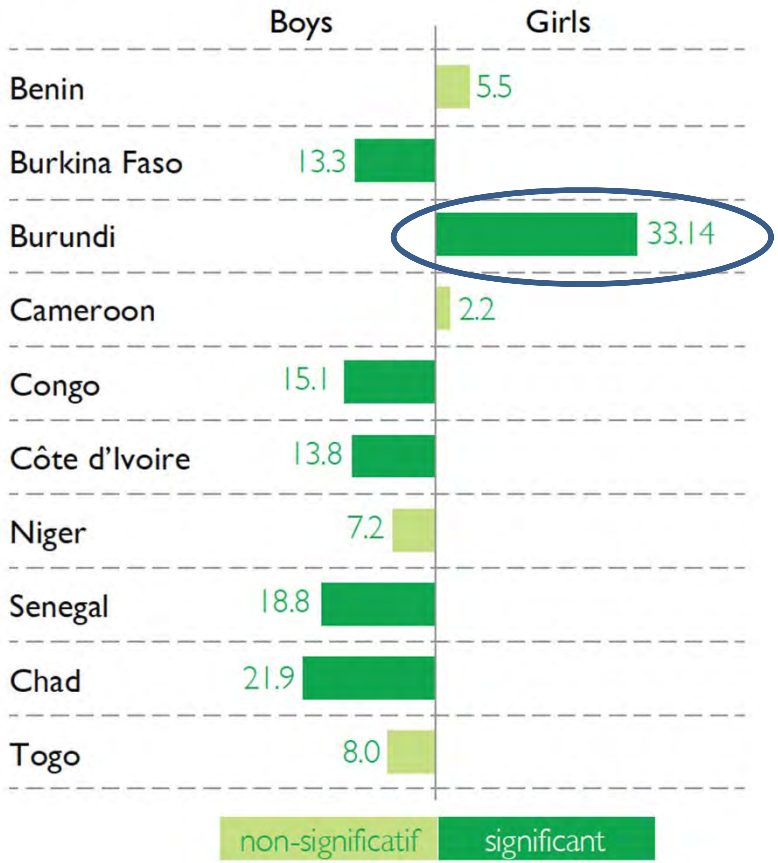
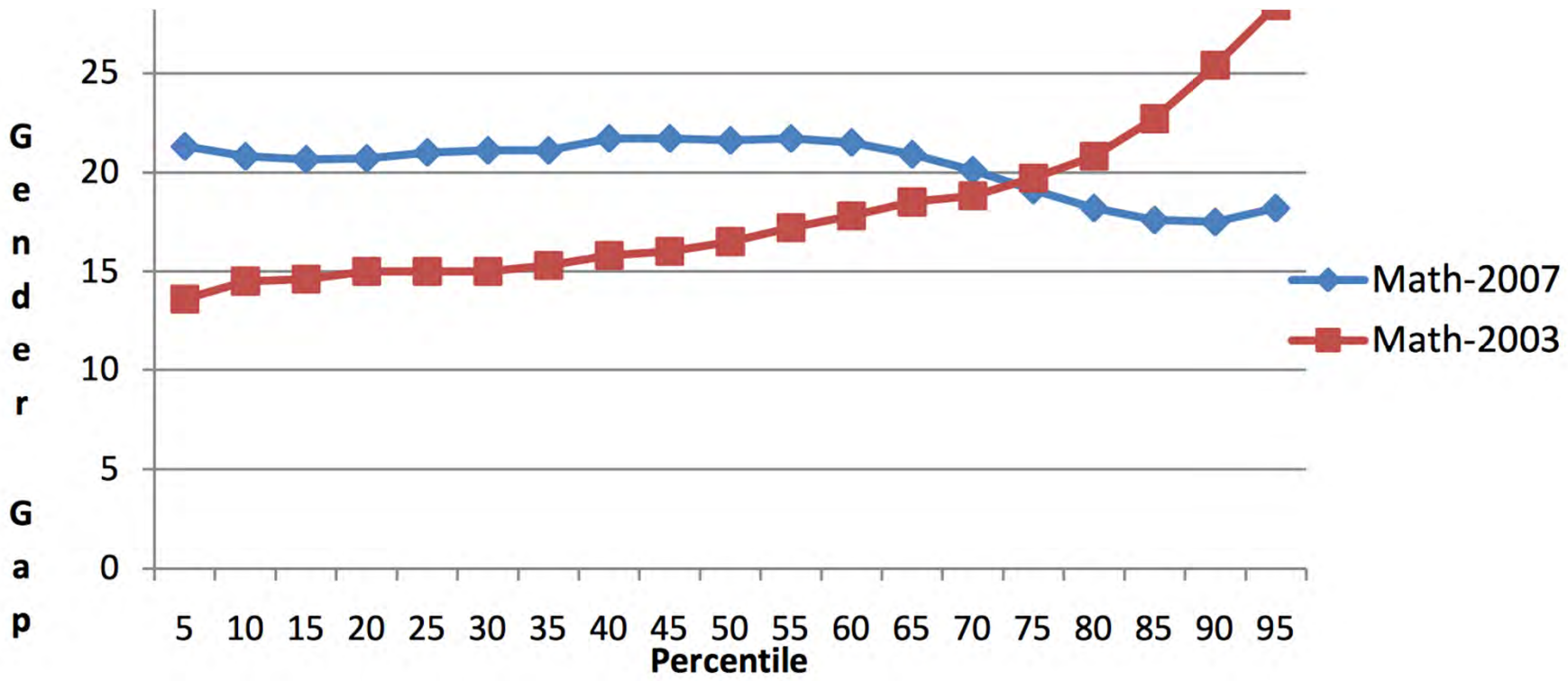


Figure 3.6: *Mathematics Performance Gap between Girls and Boys – Late Primary*



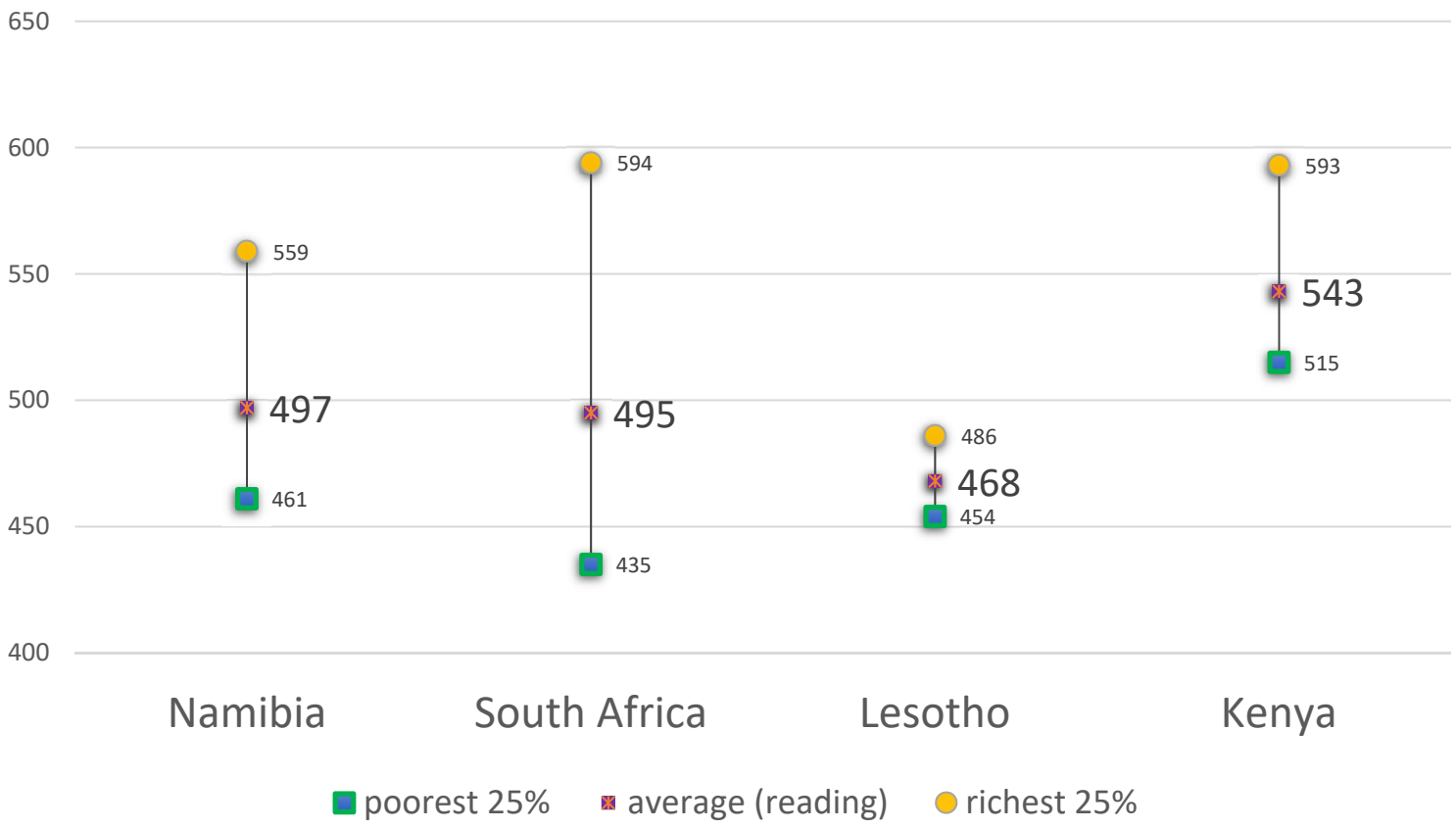
Source: PASEC 2014 Results report, 2015

Trends over time such as narrowing gender gaps amongst high performers in Ghana



Source: Sakellarion, 2012

Varying disparities in performance by socio-economic status (SACMEQ III)



Source: Spaull, 2013

Children who are bullied perform significantly worse than those who are not across all subjects in Botswana

Frequency of bullying	n	%	Mathematics			Science		
			Mean (SE)	Diff	SD	Mean (SE)	SD	Diff
At least once a month	1003	27.01	409.27(4.54)	1,2:-12.10	89.77	352.01(6.75)	133.03	1,2:-16.88
A few times a year	1953	53.93	421.37(4.28)	1,3:-39.76*	87.63	368.89(6.58)	130.34	1,3:-61.46*
Never	704	19.06	449.03(5.91)	2,3:-27.66*	81.77	413.47(8.35)	115.56	2,3:-44.58*

** Statistically significant at 5% level*

Source: Botswana Examinations Council, 2014

Evidence to understand learning outcomes relative to socio-economic and demographic conditions as well as practices or experiences

- SACMEQ data allows to look at learning achievement by sex, region, school location, socio-economic status

- PASEC by sex, school location, type and environment (teaching resources, health and hygiene, infrastructure); teacher qualifications, parental literacy and home environment, disability, working practices outside school

What can the data analysis / evidence potentially do?

- Raise awareness on key issues
- Identify priority issues for reform
- Add depth and perspective to the analysis of education systems
- Stir public debate, and place educational issues on the policy agenda
- Inform further focused studies

How can such evidence be useful for improving teaching and learning?

System	Curriculum	School	Teacher	Home
• Allocating resources • Implementing programatic reforms • outlining goals for curricular achievement	• Curriculum development • Curricular content and methods • Curriculum design	• Setting faculty priorities • Improving student support services • Enriching school environment	• Securing resources for professional development or improving pedagogical practices • Revising courses and assignments	• Inform programmes to encourage parental involvement

Direct impact of assessments on curriculum and teacher training—the case of Jordan

- Curriculum
  - Jordanian "Education Reform for Knowledge Economy/ERfKE, 2003-2009" revision of the mathematics, science and Arabic curricula
  - new textbooks developed for mathematics, science and Arabic using previously used and released PISA and TIMSS test items

- Teacher Training
  - remedial instructional strategies to support teachers' practices to enhance students' skills dealing with real-life problems
  - teachers' guides developed with questions similar to PISA and TIMSS assessment items

Source: Ababneh, Tweissi & Abulibdeh, 2016

Examples from around the world: Impact of assessments on teachers and curriculum (TIMSS & PISA)

- Canada: **development of instructional materials** based on students' responses to TIMSS test items
- Romania: **new subdomains** added to the mathematics curriculum
- Spain: **new contents** relating to real-life situations added to the mathematics and science curriculum
- Ireland: contributed to the **development** of the secondary-school science **curriculum**
- Slovak Republic: TIMSS findings affected curriculum document studies, **teacher studies**, mathematics and science **teaching methodologies**

(Source: Robitaille, Beaton and Plomp, 2000, as cited in Greaney and Kellaghan, 2008)

What about the Early Grade Reading Assessment (EGRA) and the Early Grade Math Assessment (EGMA) and the citizen-led assessments (Uwezo, Jangandoo, Beekunko, LearNigeria)?

Countries that conducted **at least one oral assessment** in Sub-Saharan Africa since 2009

Countries that conducted **at least one national oral assessment** in Sub-Saharan Africa since 2009

Source: EGRA tracker, UIS Catalogue and Database on Learning Assessments

Why so much focus on oral assessments in Sub-Saharan Africa?

- Less resource heavy
- Provide timely access to data to inform decision making improvement of reading skills
- Allow for detection of reading weaknesses early on and of children 'at risk' of dropping out
- Measure reading skills to children who are learning to read, and have not yet mastered the necessary skills to take traditional written tests
- Fill a gap in data on learning achievement for countries that do not participate in cross-national initiatives by establishing national reading and mathematics performance measures.

Limitations: Not comparable across countries, small samples, narrow focus in language/math domain, limited data disaggregation possibilities

DISCUSSION

Do you have any experiences you would like to share about the use (or lack) of data from a large-scale learning assessment – whether from PASEC, SACMEQ, EGRA/EGMA or other national assessment towards improving teaching and learning strategies?

Avez-vous des expériences que vous aimeriez partager sur l'utilisation (ou le manque) de données d'une évaluation d'apprentissage à grande échelle - du PASEC, du SACMEQ, de l'EGRA / EGMA ou d'autres évaluations nationales visant à améliorer les stratégies d'enseignement et d'apprentissage?

Limitations of large-scale learning assessments for teaching and learning

Limitations of large-scale learning assessments for teaching and learning

Too narrowly focused: do not adequately assess the broad range of competencies nor the breadth of knowledge within a given domain

LIMITED DOMAINS & COMPETENCIES

Arts
Music
Citizenship
Environmental responsibility

… are rarely assessed

WITHIN A DOMAIN
e.g. Language

Literature
Writing Skills
Listening Comprehension

… are often disregarded/not measured

Limitations of large-scale learning assessments for teaching and learning

Cannot respond to the range of expectations from diverse stakeholders – governments, schools, teachers, partners and the students themselves

For **schools** to use the assessment results to improve the school environment, they have to also be given the **data at school level** on perception of children, but also teachers and potentially parents (depending on the additional information collection)

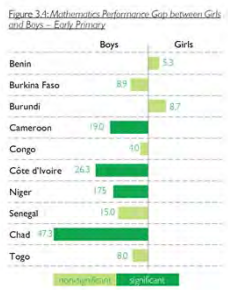
For **teachers** to improve their practice with some disadvantaged students, scores would have to be given at to them at both the **item level and the specific student level**, which happens rarely [items are not public, protect anonymity of students]

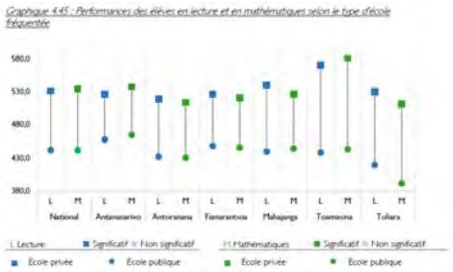
For **students** to benefit from their experience, they have to be given individual feedback on their performance, which is unlikely as data is analyzed at sub-group level

Limitations of large-scale learning assessments for teaching and learning

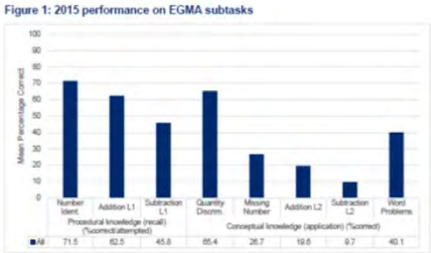
Data analyses are not generally presented in ways that enable targeting interventions at classroom level – tendencies at national level not necessarily applicable to classroom context

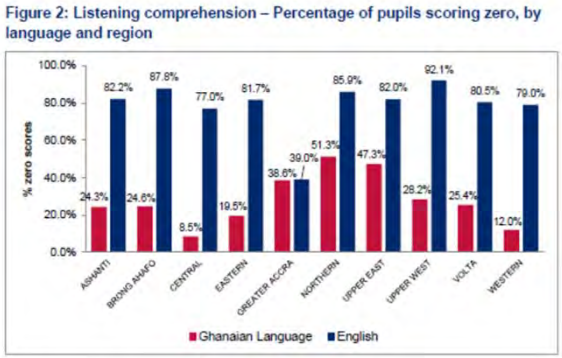
From a cross-national initiative





From EGRA/EGMA assessments





Despite limitations, there is increased pressure to improve ranks on cross-national initiatives' league tables…

Grade four		Grade eight	
Country	Average score	Country	Average score
TIMSS scale average	500	TIMSS scale average	500
Singapore	587	Singapore	567
Chinese Taipei	557	Chinese Taipei	561
Hong Kong SAR[1]	554	Japan	554
Japan	548	Korea, Rep. of	553
Russian Federation	546	England[3]	542
Latvia[2]	542	Hungary	539
England	542	Czech Republic	539
United States[3,4]	**539**	Slovenia	538
Hungary	536	Hong Kong SAR[1,3]	530
Italy	535	Russian Federation	530
Kazakhstan[2]	533	**United States[3,4]**	**520**
Germany	528	Lithuania[2]	519
Australia	527	Australia	515
Slovak Republic	526	Sweden	511
Austria	526	Scotland[3]	496
Sweden	525	Italy	495
Netherlands[5]	523	Armenia	488
Slovenia	518	Norway	487
Denmark[3]	517	Ukraine	485
Czech Republic	515	Jordan	482
Lithuania[2]	514	Malaysia	471
New Zealand	504	Thailand	471
Scotland[3]	500	Serbia[2,4]	470
Armenia	484	Bulgaria[7]	470
Norway	477	Israel[7]	468
Ukraine	474	Bahrain	467
Iran, Islamic Rep. of	436	Bosnia and Herzegovina	466
Georgia[2]	418	Romania	462
Colombia	400	Iran, Islamic Rep. of	459
El Salvador	390	Malta	457
Algeria	354	Turkey	454
Kuwait[6]	348	Syrian Arab Republic	452
Tunisia	318	Cyprus	452
Morocco	297	Tunisia	445
Qatar	294	Indonesia	427
Yemen	197	Oman	423
		Georgia[2]	421
		Kuwait[6]	418
		Colombia	417
		Lebanon	414
		Egypt	408
		Algeria	408
		Palestinian Nat'l Auth.	404
		Saudi Arabia	403
		El Salvador	387
		Botswana	355
		Qatar	319
		Ghana	303

Grade four		Grade eight	
Country	Average score	Country	Average score
TIMSS scale average	500	TIMSS scale average	500
Hong Kong SAR[1]	607	Chinese Taipei	598
Singapore	599	Korea, Rep. of	597
Chinese Taipei	576	Singapore	593
Japan	568	Hong Kong SAR[1,4]	572
Kazakhstan[2]	549	Japan	570
Russian Federation	544	Hungary	517
England	541	England[4]	513
Latvia[2]	537	Russian Federation	512
Netherlands[3]	535	**United States[4,5]**	**508**
Lithuania[2]	530	Lithuania[2]	506
United States[4,5]	**529**	Czech Republic	504
Germany	525	Slovenia	501
Denmark[4]	523	Armenia	499
Australia	516	Australia	496
Hungary	510	Sweden	491
Italy	507	Malta	488
Austria	505	Scotland[4]	487
Sweden	503	Serbia[2,5]	486
Slovenia	502	Italy	480
Armenia	500	Malaysia	474
Slovak Republic	496	Norway	469
Scotland[4]	494	Cyprus	465
New Zealand	492	Bulgaria	464
Czech Republic	486	Israel[7]	463
Norway	473	Ukraine	462
Ukraine	469	Romania	461
Georgia[2]	438	Bosnia and Herzegovina	456
Iran, Islamic Rep. of	402	Lebanon	449
Algeria	378	Thailand	441
Colombia	355	Turkey	432
Morocco	341	Jordan	427
El Salvador	330	Tunisia	420
Tunisia	327	Georgia[2]	410
Kuwait[6]	316	Iran, Islamic Rep. of	403
Qatar	296	Bahrain	398
Yemen	224	Indonesia	397
		Syrian Arab Republic	395
		Egypt	391
		Algeria	387
		Colombia	380
		Oman	372
		Palestinian Nat'l Auth.	367
		Botswana	364
		Kuwait[6]	354
		El Salvador	340
		Saudi Arabia	329
		Ghana	309
		Qatar	307

…resulting in large-scale learning assessments becoming more **high-stakes**

- For governments
- For schools
- For teachers

Evidence shows that **high-stakes examinations** can lead to

- Narrowing curricular content to areas that can be measured
- Teaching to the test
- Neglecting a broad range of competencies for personal development

DISCUSSION

What are your thoughts on whether these unintended consequences are also applicable to large-scale assessments, which generally are administered to a small sample of students?

Que pensez-vous de ces conséquences imprévues? Est-ce qu’elles s'appliquent également aux évaluations à grande échelle, qui sont généralement administrées à un petit échantillon d'élèves?

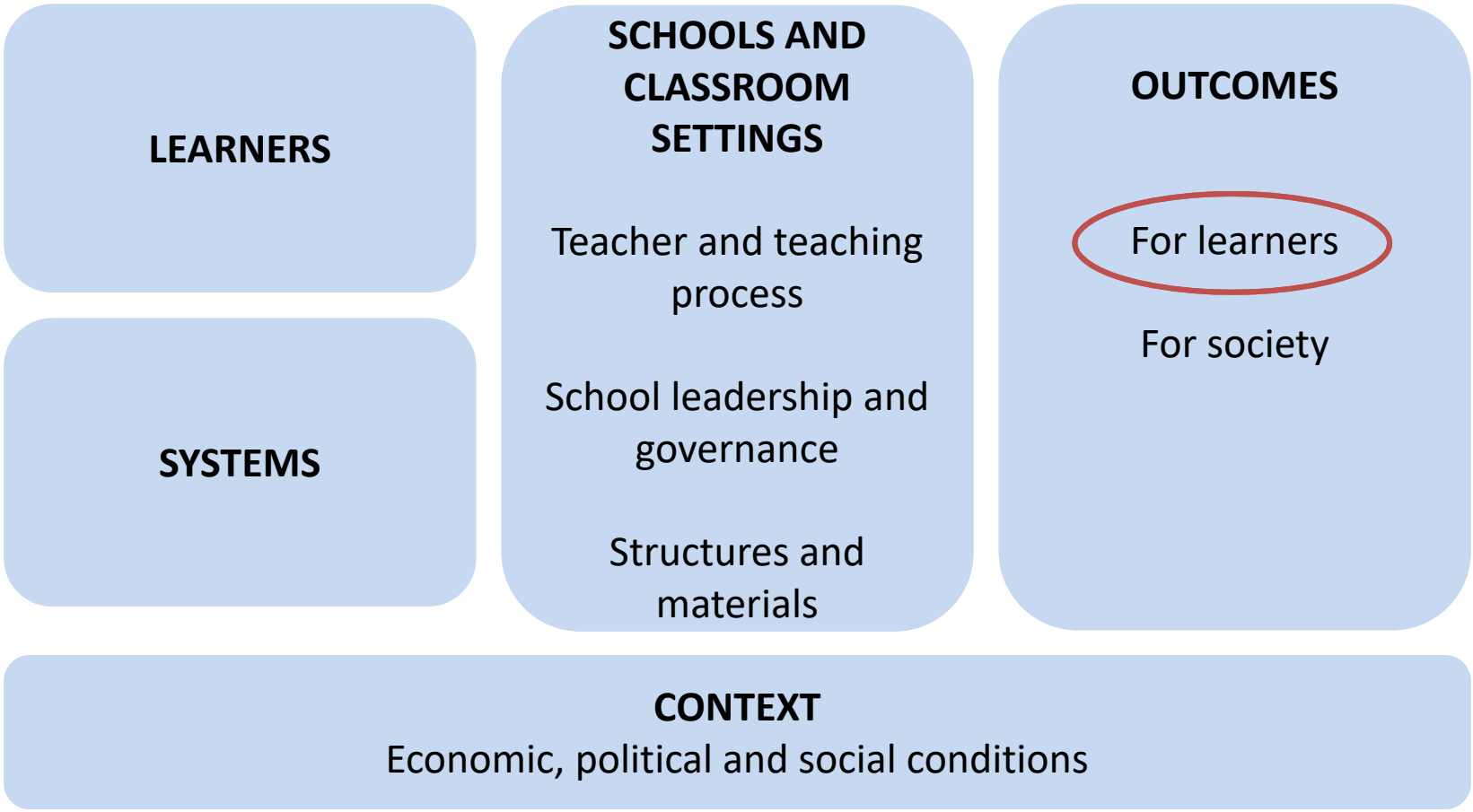
If large-scale learning assessments cannot be used to **directly** influence teaching and learning, can school-based assessment results be aggregated to provide a system-level measurement of learning?

DISCUSSION

It may seem feasible and cost-effective; are there any limitations you can think of?

Cela peut sembler faisable et rentable; Y a-t-il des limites auxquelles vous pouvez penser?

A Framework for quality education



Source: UNESCO GEM report 2016 (adapted from GMR 2005)

Thank you

Learn more: www.unesco.org/education

@UNESCO

Maya Prince and Zi Hu
Partnerships, Cooperation and Research
Education 2030
en.unesco.org/themes/education-21st-century
https://en.unesco.org/themes/learning-assessment