

Policy Brief

What is the Value of Internationally Comparative Learning Assessment Programmes in Africa?



Association for the
Development
of Education
in Africa

Introduction

Education has long been characterized as key to national economic development and political democratization in the contemporary world. The global discourse on post 2015 education goals has increasingly focused on the quality of education and learner achievement. This is likely to give added impetus to the rapid growth in the use of International Comparative Learner Assessments (ICLAs) by countries. The main reason for this growing trend is the shared opinion that countries will need to be able to compete in the knowledge economy to assure the economic wellbeing of their citizens. Whilst benchmark indicators of knowledge economy ‘supply’ variables, such as investment in education as a proportion of GDP, have been available for a long time, countries had no way of comparing the effect of their investments and schooling in general upon learners’ knowledge and skills.

International benchmarking has been identified as the “basis for improvement. ... It is only through such benchmarking that countries can understand relative strengths and weaknesses of their education systems and identify best practices and ways forward” (OECD 2006, 18). International consensus has emerged—at least among “developed” countries—about the legitimacy and, even more so, the necessity of international testing and national assessment. Moreover, an expanding number of donor agencies and multilateral organizations are mandating some form of learning assessment to accompany their loans and other aid support. Often involving cooperation with organizations that provide technical advice and support, testing is increasingly viewed as an obligation of countries¹.

Challenges

Constraints on testing and assessment exist. There is a fundamental lack of resources and lack of organizational capacity to administer this complex activity. As more standardization of testing regimes is required, more administrative capability and expertise are also needed to carry out related activities. Sampling designs, test construction, translation and so on all become complex and important issues. Poor countries simply may not have the necessary resources to coordinate and undertake these activities. Thus, to date, many African countries have been absent from the populations of countries in

which international testing has occurred. Second, engaging in testing opens a country to external scrutiny and may also intensify internal conflicts over perceived responsibilities for inadequate or unequal educational outcomes. Even the most progressive countries may feel that such comparisons are premature and dangerous at a time when they may be trying to devise school systems and curricula that bridge gaps among groups in their countries.

The usefulness of international comparative tests is that they are being increasingly used to inform education policy and governance. However, issues are raised as to whether single tests can give comparable measures of curriculum effects across countries. Research suggests that the spectre of an “idealized international curriculum” lies behind even the most sophisticated research designs and given the high costs of participation for countries, both politically and financially, it is important that African policy makers are informed as to the value of engaging in such processes. It is also imperative that the African Union understands the implications and challenges associated with using and advancing international comparative scores on learning outcomes for its member states as a way of benchmarking progress.

How do international comparative assessments differ from regional testings?

(PIRLS) AND TIMMS are a series of international assessments of reading achievement and science and mathematics among fourth graders. It is conducted by the International Association for the Evaluation of Educational Achievement (IEA). For the first time in the 2011 cycle, prePIRLS was offered to assess basic reading skills as a bridge to PIRLS, for countries where most children are still developing fundamental reading skills at the end of the primary school cycle. TIMSS is also conducted in Grade 8, which is considered lower secondary by some countries and primary by others.

Another worldwide study that evaluates students’ math and science performance is the Programme for International Student Assessment (PISA) and is conducted by OECD. The PISA measures advanced mathematical, scientific, and reading literacy and problem solving skills and is primarily administered in high-income and OECD countries. PISA also includes

¹ Globalization and the Growth of International Educational Testing and National Assessment David H. Kamens and Connie L. McNeely Electronically published December 8, 2009



measures of general or cross-curricular competencies, such as problem solving and computer literacy. PISA aimed at 15 year olds emphasizes functional skills that students have acquired as they near the end of compulsory schooling.

TIMSS and PISA are not the same. TIMSS is curriculum-based, reflecting the skills and knowledge taught in schools. PISA assesses whether students can apply what they've learned to solve "real world" problems. PISA tests an age-based sample (15 year olds). PIRLS and TIMSS are grade-based (4th and 8th graders). PISA is overseen by representatives from participating governments who meet under the auspices of the Organization for Economic Cooperation and Development (OECD). TIMSS and PIRLS are governed by a consortium of researchers and government representatives known as the International Association for the Evaluation of Educational Achievement (IEA).

The simpler assessments may be viewed as measuring "minimum standards" in literacy and numeracy. ASER and Uwezo use the same Primary Level 2 curriculum as the basis for the testing materials used for all children age 6- 16 regardless of their enrolment status or grade level. While it is important to know, for example, that only 71 percent of students in Level 6 can read a Level 2 story, ASER and Uwezo do not provide data on how many Level 6 students are reading and comprehending on grade level.

The Early Grade Reading Assessment (EGRA) is an oral student assessment designed to measure the most basic foundation skills for literacy acquisition in the early grades: recognizing letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension. EGRA methodology was applied to 11 countries and 19 languages. It has been adopted and used by other implementing partners in more than 30 other countries and more than 60 other languages.



African Union initiative

The AU EMIS Restricted Technical Committee at its annual meeting in April 2014 in reviewing the contributions from members, including SACMEQ which is a member, noted that it was pertinent to its work that it engages more systematically in the role of international comparative learning assessments in Africa.

It was noted that in order to deliver on its mandate in monitoring of country performance on the African Union's Plan of Action, there is a need to continentally monitor learning outcomes across all member states, particularly at the school level. There is not a single comparable system producing comparable learner achievement statistics. Some countries use PASEC, others SAQMEC, more recently PISA, PIRLS and TIMSS -all of whom use different test items and methodologies, target groups and timelines.

The post-2015 agenda focus on quality of education invariably means an emphasis on the use of learning assessments. International learning assessment systems are different to national systems and their implications are not fully understood by countries. Country experiences participating in ICLAs could inform better practices for the continent. The costs and benefits need closer review.

Internationally Comparative Learning Assessment Programmes used in Africa

In Africa there are a number of internationally comparative tests at various school levels. The majority of international learning assessments are conducted during the primary grades. There are two distinct approaches to assessment of foundation skills during this time—the smaller, quicker, cheaper assessments such as ASER, EGRA, EGMA, Literacy Boost, and Uwezo² - and large-scale assessments such as PASEC, PISA, PIRLS, SACMEQ, and TIMSS. Each approach has advantages and limitations

The most well know include the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ), started in 1990, which has 15 member countries; seven of whom participated in its first assessment study in 1995–97, and 14 were involved in 2000–2003 studies and 16 in the 2006–2011 studies. Similarly, Francophone Africa has the Conference des Ministres de l'Éducation des Pays Ayant le Français en Partage (CONFEMEN's PASEC), which encourages and assists member states in performing national assessments. Assessments such as PASEC, and SACMEQ are based

² See Appendix for details on these tests.

on standards derived from national curriculums across the region. The content is the result of consensus building across ministries of education on what students at a given grade level should know and be able to do. An added advantage of the PASEC and SACMEQ assessments is that they measure learning at the beginning and end of the academic year, allowing countries to measure progress of the same students over time. They target different learner groups and have a different scope³.

Current Initiatives to review ICLAs

Learning Metrics Task Force hosted by the Brookings Institute and representing 20 organizations, was created in 2013 to shift in the global education conversation from access to access plus learning, and to build consensus on global learning indicators and actions to improve the measurement of learning in all countries. With a common commitment to a highly consultative process, the initiative was structured around three guiding questions:

- What learning is important globally?
- How should it be measured?
- How can measurement of learning improve education quality?

Emerging from this collaboration, some key initiatives are emerging to examine ways to align and synthesize existing efforts.

The UNESCO Institute for Statistics (UIS) has launched the Observatory of Learning Outcomes to track learning achievement and measurement worldwide. The Observatory will rely on partnerships with research institutions that are collecting data on learning and synthesize the data in a central database. While the UIS will not create a new assessment, it will promote convergence among existing initiatives. In order to increase the body of cross-national comparable data that is available.

The Learning Champion Framework, co-chaired by ADEA and the Aga Khan Foundation, was launched in October 2014 to provide an overview of the in country-level process to: convene stakeholders; diagnose the quality of their assessment systems; and assess the necessary technical and financial resources required to improve learning measurement and outcomes. At the end of the Learning Champion initiative, the goal is to have achieved the following concrete outcomes across all countries: a set of country-level validated tools and strategies that permit education systems to measure learning accurately and verifiably across seven learning domains.

There are also efforts underway for some of the large-scale assessment bodies to align their assessment efforts. In 2011, the UNESCO International Institute for Educational Planning (IIEP) hosted the second exchange meeting between PASEC and SACMEQ with the goal of creating a common vision for basic education in Africa. SACMEQ has held conversations with IEA regarding including some of the TIMSS and PIRLS assessment items within the SACMEQ assessment for comparability. Furthermore, SACMEQ, PASEC, and LLECE, in collaboration with UIS intended to establish a partnership to discuss the feasibility of aligning and/or comparing efforts.

3 PASEC focuses on Grades 2 and 5 and tests Math, Reading (French). Some one to three countries are annually tested and the assessments range from 1993-2010. Some 13 African countries participate. SACMEQ focuses on Grade 6 and tests Math, Reading (English) and assessments are conducted in 1995, 2000, and 2007. Some 15 African countries participate.

UNESCO is working with an interagency team to develop the Holistic Early Childhood Development Index (HECDI) to monitor global progress toward EFA Goal 1. The final version of this tool will contain child outcome indicators that may include early development and/or learning.

Critical Questions for consideration

The growing use of international comparative tests projects the world as a single space of comparative and equal measurement of the performance of school systems. They produce league tables on countries' education performance. National school systems are variously positioned within the global market place and the global educational policy field is shaped by the discourse on these positions with important effects on national policy-making⁴. Are testing regimes becoming the major policy steering systems and the work of schools today?



What are the implications for Africa? How do comparable learning statistics help countries make decisions on improving the quality of their education systems?

- The international testing movement itself rests on the belief that one can discover methods to increase student achievement in different subjects and that these techniques are valid across all educational systems. Moreover, on the basis of such beliefs in testing, officials and educators express optimism that progress toward national educational goals can be measured and that scientific understanding can inform policy making and practices⁵. How do these tests make countries take actions to improve the learning achievement of their students?
- Research indicates that if you remove the outlying countries (the extreme performers) on comparative learning achievement tests, the level of correlation of results produced by multiple international comparative tests on countries is likely to be 60 per cent. How accurate are ICLAs in indicating common causes of country differences in learning achievement, such as educational quality, culture, wealth or a homogenous underlying factor of cognitive competence?
- Is it useful to have a continentally harmonized and comparable learner achievement statistic for primary and secondary levels? Should learner achievement scores remain exclusive to countries and not be internationally comparable given their national differentiations?
- How can ICLA players at school level harmonise their methodologies and target groups?
- What can be done about the missing African countries who do not participate in ICLAs? What should the recommendation to the African Union on urging countries to adopt this practice?

4 Jo-Anne Baird, Talia Isaacs, Sandra Johnson, Gordon Stobart, Guoxing Yu, Terra Sprague & Richard Daugherty. (2011) POLICY EFFECTS OF PISA. Oxford University Centre for Educational Assessment. UK
5 David H. Kamens and Connie L. McNeely (2009). Globalization and the Growth of International Educational Testing and National Assessment Electronically published December 8,

ICLA-Programmes in Africa

Organizations	Name	Grade level	Type of test			Remarks
			Reading	Maths	Science	
EDI	Early Development Instrument	Just before entry to grade 1				Readiness test based on the criteria: 1. Physical health & well-being, 2. Emotional maturity, Social competence' 3. Language & cognitive development, 4. Communication skills & general knowledge
EGRA	Early Grade Reading Assessment	Early grades	X			The Early Grade Reading Assessment (EGRA) is an oral learner assessment designed to measure the most basic foundation skills for literacy acquisition in the early grades: recognizing letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension.
EGMA	Early Grade Math Assessment	Early grades		X		EdData II developed the EGRA methodology and has applied it in 11 countries and 19 languages. It has been adopted and used by other implementing partners in more than 30 other countries and more than 60 other languages. Data from EGRA have been used for feedback on teacher practice in rigorous but easy-to-understand ways. Many countries have shown an interest in using it as a springboard to improve reading, and have gone on to redesign their teacher training around reading.
SACMEQ *	The Southern and Eastern Africa Consortium for Monitoring Educational Quality	Grade 6	X	X		SACMEQ III: Knowledge about HIV and AIDS, Gender Equality, School Inputs, Trends in Pupil Achievement
PASEC *	Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC)	2, 5				The aim is to link the acquisition means learners in French, mathematics and sometimes national language with the material and organizational conditions in which they are enrolled levels. For this, however, cannot be limited to mean scores of learners at the end of the school year because all the past learners would intervene in this action and is not really what is not attributable to the year assessed against previous years. This involves having a measure of the level of learner achievement at the beginning and end of the school year, which is a high specificity of the methodology PASEC. There is talk of value-added models.
Uwezo *		Early grades (2)	X	X	-	Uwezo means 'capability' in Kiswahili. Uwezo is a four year initiative that aims to improve competencies in literacy and numeracy among children aged 6-16 years old in Kenya, Tanzania and Uganda, by using an innovative approach to social change that is citizen driven and accountable to the public.
PISA *	Program for International Student Assessment	15 years old	X	X	X	PISA tests critical thinking in math, science, and reading to 15 year olds. The test questions do not measure memorization of facts, but rather demand that learners draw on knowledge and real-world problem solving skills
prePIRLS	PrePIRLS	4,5,6				prePIRLS is a stepping stone to participating in PIRLS and provides a way to assess reading at the end of the primary school cycle for a range of developing countries. prePIRLS reflects the same conception of reading as PIRLS, except it is less difficult. Depending on a country's educational development, prePIRLS can be given at the fourth, fifth, or sixth grade.
PIRLS	Program for International Reading Literacy Study	Grade 4				The IEA Progress in International Reading Literacy Study (PIRLS) 2011 was the third assessment, after PIRLS 2001 and PIRLS 2006 , in a five-year cycle of assessments measuring trends in reading literacy achievement in primary school.
TIMSS		4 and 8		X	X	TIMSS are a series of international assessments of achievement and science and mathematics in fourth and eighth grade. They are curriculum based.
EGRA	Early grade Reading Assessment	Early grades	X			The Early Grade Reading Assessment (EGRA) is an oral learner assessment designed to measure the most basic foundation skills for literacy acquisition in the early grades: recognizing letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension.
EGMA	Early grade Math Assessment	Early grades		X		EdData II developed the EGRA methodology and has applied it in 11 countries and 19 languages. It has been adopted and used by other implementing partners in more than 30 other countries and more than 60 other languages. Data from EGRA have been used for feedback on teacher practice in rigorous but easy-to-understand ways. Many countries have shown an interest in using it as a springboard to improve reading, and have gone on to redesign their teacher training around reading
SACMEQ *	The Southern and Eastern Africa Consortium for Monitoring Educational Quality	Grade 6	X	X		SACMEQ III: Knowledge about HIV and AIDS, Gender Equality, School Inputs, Trends in Pupil Achievement
PASEC *	Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC)	2, 5				The aim is to link the acquisition means learners in French, mathematics and sometimes national language with the material and organizational conditions in which they are enrolled levels. For this, however, cannot be limited to mean scores of learners at the end of the school year because all the past learners would intervene in this action and is not really what is not attributable to the year assessed against previous years. This involves having a measure of the level of learner achievement at the beginning and end of the school year, which is a high specificity of the methodology PASEC. There is talk of value-added models.
Uwezo *		Early grades (2)	X	X		Uwezo means 'capability' in Kiswahili. Uwezo is a four year initiative that aims to improve competencies in literacy and numeracy among children aged 6-16 years old in Kenya, Tanzania and Uganda, by using an innovative approach to social change that is citizen driven and accountable to the public.
PISA *	Program for International Student Assessment	15 years old	X	X	X	PISA tests critical thinking in math, science, and reading to 15 year olds. The test questions do not measure memorization of facts, but rather demand that learners draw on knowledge and real-world problem solving skills
prePIRLS	PrePIRLS	4,5,6				prePIRLS is a stepping stone to participating in PIRLS and provides a way to assess reading at the end of the primary school cycle for a range of developing countries. prePIRLS reflects the same conception of reading as PIRLS, except it is less difficult. Depending on a country's educational development, prePIRLS can be given at the fourth, fifth, or sixth grade.
PIRLS	Program for International Reading Literacy Study	Grade 4				The IEA Progress in International Reading Literacy Study (PIRLS) 2011 was the third assessment, after PIRLS 2001 and PIRLS 2006 , in a five-year cycle of assessments measuring trends in reading literacy achievement in primary school.
TIMSS		4 and 8		X	X	TIMSS are a series of international assessments of achievement and science and mathematics in fourth and eighth grade. They are curriculum based.

Table 1: Mathematics Achievement		
Country	Average Scale Score	SE
Korea, Rep. of	613	2.9
Singapore	611	3.8
Chinese Taipei	609	3.2
Hong Kong SAR	586	3.8
Japan	570	2.6
Russian Federation	539	3.6
Israel	516	4.1
Finland	514	2.5
United States	509	2.6
England	507	5.5
Hungary	505	3.5
Australia	505	5.1
Slovenia	505	2.2
Lithuania	502	2.5
TIMSS Scale Centrepoint 500		
Italy	498	2.4
New Zealand	488	5.5
Kazakhstan	487	4
Sweden	484	1.9
Ukraine	479	3.9
Norway	475	2.4
Armenia	467	2.7
Romania	458	4
United Arab Emirates	456	2.1
Turkey	452	3.9
Lebanon	449	3.7
Malaysia	440	5.4
Georgia	431	3.8
Thailand	427	4.3
Macedonia Rep. of	426	5.7
Tunisia	425	2.8
Chile	416	2.6
Iran, Islamic Rep. of	415	4.3
Qatar	410	3.1
Bahrain	409	2
Jordan	406	3.7
Palestinian Nat. Auth.	404	3.5
Saudia Arabia	394	4.6
Indonesia	386	4.3
Syrian Arab Rep.	380	4.5
Morocco	371	2
Oman	366	2.8
Ghana	331	4.3
Ninth Grade Participants		
Botswana	397	2.5
South Africa	352	2.5
Honduras	338	3.7

Table 2: Science Achievement		
Country	Average Scale Score	SE
Singapore	590	4.3
Chinese Taipei	564	2.3
Korea, Rep. of	560	2
Japan	558	2.4
Finland	552	2.5
Slovenia	543	2.7
Russian Federation	542	3.2
Hong Kong SAR	535	3.4
England	533	4.9
United States	525	2.6
Hungary	522	3.1
Australia	519	4.8
Israel	516	4
Lithuania	514	2.6
New Zealand	512	4.6
Sweden	509	2.5
Italy	501	2.5
Ukraine	501	3.4
TIMSS Scale Centrepoint 500		
Norway	494	2.6
Kazakhstan	490	4.3
Turkey	483	3.4
Iran, Islamic Rep. of	474	4
Romania	465	3.5
United Arab Emirates	465	2.4
Chile	461	2.5
Bahrain	452	2
Thailand	451	3.9
Jordan	449	4
Tunisia	439	2.5
Armenia	437	3.1
Saudia Arabia	436	3.9
Malaysia	426	6.3
Syrian Arab Rep.	426	3.9
Palestinian Nat. Auth.	420	3.2
Georgia	420	3
Oman	420	3.2
Qatar	419	3.4
Macedonia Rep. of	407	5.4
Lebanon	406	4.9
Indonesia	406	4.5
Morocco	376	2.2
Ghana	306	5.2
Ninth Grade Participants		
Botswana	404	3.6
Honduras	369	4
South Africa	332	3.7

