DRAFT - FOR CONSULTATION

Learning Metrics Task Force

Discussion Paper #1: Multi-Country Assessments of Learning

Kate Anderson Simons, Brookings Institution, Center for Universal Education

Please send feedback to metrics@globalcompactonlearning.org

This discussion paper is the first in a series of discussion papers to inform the work of the *Learning Metrics Task Force*. The *Learning Metrics Task Force* (LMTF) was convened to make recommendations for learning goals and targets regarding early, primary, and post-primary learning. These recommendations will be designed for use at the global level, building off the ongoing work on measuring learning that is happening at country and regional levels. The work of the task force is guided by a secretariat (UNESCO Institute for Statistics and Brookings Institution, Center for Universal Education) and three co-chairs from the civil society, multilateral, and private sectors. Building on the advances in access to education, the desired impact of this project is to ensure that learning features prominently as a central objective of education development, particularly within low- and middle-income countries.

Throughout the duration of the initiative, the task force will engage stakeholders in a consultative process involving in-person and on-line consultations, interviews, and best practice review. In preparation for the meetings and consultations, the secretariat, in collaboration with the working groups and task force, will prepare a series of research papers to inform the task force deliberations. The documents will be made available for public comment. The publications will include:

- **Multi-Country Assessments of Learning:** This paper will describe existing initiatives to measure learning at the national, regional, and international levels.
- **National Assessments of Learning:** This paper will examine country policies on learning; current methods for measurement of learning in countries; and the state of community movements for learning.
- **Global Debates in Learning Assessment:** This paper will analyze the issues that must be deliberated by the *Learning Metrics Task Force*, including different philosophies and approaches to measuring learning.
- **Equity Dimensions of Learning Assessment** This paper will explore the research on accessibility, language of administration, and disaggregation of data.
- **Good Practices Review from the Field:** This desk review of learning evaluation reports and survey of country officers will compile good practices in measurement and application of learning outcomes.

Special thanks to Cesar Guadalupe and Maya Prince of the UNESCO Institute for Statistics and Rebecca Winthrop, Anda Adams, Justin van Fleet, and Nana Debrah-Karikari of the Brookings Institution Center for Universal Education for their invaluable contributions to this paper.

Learning Metrics Task Force

Discussion Paper #1: Multi-Country Assessments of Learning – DRAFT 29 JUNE 2012

Kate Anderson Simons, Center for Universal Education at the Brookings Institution

Introduction and purpose

The Education for All (EFA) goals initiated in 1990 in Jomtien, Thailand demonstrated a commitment to meeting basic learning needs. This commitment was restated in 2000 in the Dakar Framework for Action, in which Goal 6 states; "Improving every aspect of the quality of education, and ensuring their excellence so that recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy and essential life skills".[1] Yet today there is growing evidence suggesting that millions of children and youth do not have the basic skills and knowledge necessary to succeed in school and life. Since 2000, there have been extraordinary gains in access to education.[2] However, many of those who do complete the education cycle do not finish with the skills needed to fully participate in society and the economy. The true cost to society is impossible to measure within current assessment systems. While international assessments are used to compare learning across countries, they do not include the majority of the world's children living in low- and low-middle-income countries,[3] particularly the most vulnerable children and youth. Other multi-country assessments are conducted following approaches that prevent their results from being pooled into a unique set of equivalent evidence.

There are many efforts underway to assess learning at a national or international level and increasing international interest in the topic; approximately 156 countries currently measure learning levels through national assessments and examinations or participation in international, regional, or cross-national assessment initiatives.[4] There is increasing international support by UN agencies, bilateral donors, philanthropic foundations, and others for improving capacity for measuring learning and developing a shared vision for learning globally.[5] With universal primary education becoming a reality, there is an unprecedented opportunity to shift the global education agenda to "access plus learning."

This paper is intended to take stock of current multi-country efforts to collect and use data on learning. First, the characteristics of multi-country learning assessments are presented. Next, 16 instruments and initiatives that measure learning with an international, regional, or cross-national scope are profiled. The following section discusses the strengths and gaps of the current efforts. The paper concludes with a summary of upcoming efforts to improve global measurement of learning and next steps.

Characteristics of learning assessments

[1] Cesar Guadalupe, "Learning outcomes: the key challenge," Montreal: UIS, 2012.

[2] UNESCO, "EFA Global Monitoring Report," Paris: UNESCO, 2011.

[3] In this paper, developing countries are defined as the 146 countries categorized by the World Bank as low- or middle-income. Appendix A lists all developing countries considered for this paper. Source: data.worldbank.org.

[4] Guadalupe, "Learning outcomes."

[5] Nicholas Burnett, "A learning goal for education? Scoping for MDG2 post-2015: Report to Hewlett Foundation/Wellspring Advisors," Washington, DC: Results for Development Institute, 2012.

Three types of multi-country assessments were reviewed for this paper[6]:

- *International assessments*, which are conducted in various languages multiple regions throughout the world[7]
- *Regional assessments*, which are typically conducted in one language and used in a specific region, and
- *Cross-national assessments*, which are conducted in a small number of countries and often in multiple languages within countries.

Countries also use *national assessments and examinations* to inform policy and education reform efforts; however, for the purposes of this paper, we focus on assessments used in at least two countries. A subsequent discussion paper in this series will profile national-level assessments of learning.

Learning can be measured across multiple subjects and content areas, including language, literacy, mathematics, science, social and emotional development, citizenship skills, technology. Some countries set standards and assess learning in creativity, religion, and other context-specific skills and abilities. The majority of assessments reviewed for this paper focused on literacy/language and numeracy/mathematics, but there are efforts to assess across multiple domains, especially in the early childhood years.

Assessments are administered by a variety of actors and for a variety of purposes. National governments, civil society groups, international non-governmental organizations (NGOs), UN agencies, and donors all regularly administer learning assessments. The data are used to inform policy and practice through research, evaluation, and diagnostics of the situation of learning in a given country or region.

Recognizing that learning begins at birth and that the human brain begins forming shortly after conception, it is important to consider systems for learning assessment throughout the educational lifespan. While the goals and metrics are different in the early childhood, primary, and post-primary years, each of these unique periods are critical to successful outcomes in school and in life.

The assessments profiled in this report are intended to provide examples of how civil society, national governments, researchers, and international organizations have developed efforts to measure learning for a variety of purposes. This list is not exhaustive and the author welcomes suggestions of other instruments and initiatives for inclusion.

Profiles of Assessment Instruments and Initiatives

This section provides detailed information on existing efforts to measure learning at the cross-national, regional, and international levels. The following instruments and initiatives are profiled:

[6] This section contains profiles of learning assessments that are used in more than one country. The list of assessments is not exhaustive, and not all assessments that bear the same name are standardized across countries or meant for cross-country comparison.

[7] It is important to note that some international assessments are conducted primarily in two regions, North America and Western Europe.

Table 1: International, Regional, and Cross-National Initiatives to Measure Learning

Instrument/Initiative	Ages/ Grades	Subjects	Frequency	Location	Countries[8]	Administered by	Data availability
Multiple Indicator Cluster Survey (MICS)	Age 0-5	Early childhood development (literacy, numeracy, physical, social-emotional and approaches learning domains)	1995, 2000, 2005, 2009-11	Household	55 countries, 50 developing	Government organizations, with technical assistance from UNICEF	Full dataset available for download online
Young Lives	Ages 4-17	Language, literacy, numeracy, social/ emotional	2002, 2006, 2010,	Household	4 developing countries (Ethiopia, India, Peru, Vietnam)	Universities, independent research institutes, government research institutes, overseen by University of Oxford	Full dataset available for download online
Early Development Instrument (EDI)	Age 4-6 (School entry)	Physical, social, emotional, language, cognitive, communication	Since 1998, Varies by country	School	24 countries, 14 developing	Varies by country	Data from developing countries are held by governments, World Bank and Aga Khan University; some reports available upon request
Early Grade Reading Assessment (EGRA)	Grade 1-4	Basic literacy	Since 2008, Varies by country	School	44 developing countries	Varies by country (primarily RTI International)	Reports available online

[8] "Countries" refers to any country where the assessment is used, either at the national or sub-national level.

Early Grade Math Assessment (EGMA)	Grade 1-4	Basic math	Since 2011, Varies by country	School	11 developing countries	RTI International	Reports available online
Literacy Boost	Grade 1-4	Basic literacy	Since 2009, Varies by project	School	9 developing countries	Save the Children	Reports available upon request
Annual Status of Education Report (ASER)	Age 6-16	Reading, Math	Annually since 2005 in India, 2008 in Pakistan	Household	2 developing countries (India and Pakistan)	Civil society organizations	Public reports online disaggregated by district
Uwezo	Age 5-16	Reading, Math	Annually since 2010	Household	3 developing countries (Kenya, Tanzania, Uganda)	Civil society organizations	Public reports online disaggregated by district
Latin American Laboratory for Assessment (LLECE)	First study: Grade 3 & 4; Second study: Grade 3 & 6	Math, Reading, Science (second study only)	1997, 2006	School	First study: 13 developing countries in Latin America; Second study: 16 developing countries in Latin America	National governments	Data available for download online
Progress in International Reading Literacy Study (PIRLS)	Grade 4	Reading comprehension	2001, 2006, 2011	School	49 countries, 13 developing	National research partners	Data available for download online
Pre-PIRLS	Grades 4-6	Basic reading comprehension	2011	School	3 developing countries	National research partners	Data available for download online
Trends in International Mathematics and Science Study (TIMSS)	Grades 4 & 8	Math, Science	1995, 1999, 2003, 2007, 2011	School	63 countries, 28 developing	National research partners	Data available for download online

Analysis Programme of the CONFEMEN Education Systems (PASEC)	Grades 2 & 5	Math, Reading (French)	1993-2010 (1-3 countries assessed each year)	School	13 developing countries - Francophone Africa	National governments	Reports available for download online
Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)	Grade 6	Math, Reading (English)	1995, 2000, 2007	School	14 developing countries – Anglophone Africa	National governments	Reports available for download online
Programme for International Student Assessment (PISA)	Age 15	Literacy, Math, Science	2000, 2003, 2006, 2009	School	75 countries, 32 developing	International contractors, national governments	Reports available for download online
Literacy Assessment Monitoring Programme (LAMP)	Age 16-64	Literacy	Varies by country	Household	12 developing countries	National governments	Varies by country

Instrument: MICS Early Childhood Development Index (ECDI)

Initiative: UNICEF Multiple Indicator Cluster Surveys (MICS)

Type: International

Description: MICS is an international household survey program developed by UNICEF covering the areas of health, child development, education, child protection and HIV/AIDS. It is currently in its fourth round of administration (MICS4). An important feature of MICS is the ability to disaggregate the data to reveal important inequities faced by children such as those related to gender, area of residence, ethnicity and household poverty.

Learning measures: The fourth round of administration (MICS4) includes the Early Child Development Index (ECDI), a 17-item module used to calculate whether children are developmentally on-track on literacy, numeracy, physical, social-emotional and approaches learning domains. The ECDI is based on parent reporting of children's development.

MICS4 also contains items related to the environment which are associated with learning,[9] including:
Learning materials:

- Percentage of children under-five who have three or more children's books
- Percentage of children under-five with two or more playthings

Inadequate care:

- Percentage of children under-five left alone or in the care of another child younger than 10 years of age for more than one hour at least once in the past week

Support for learning:

- Percentage of children 36-59 months with whom an adult has engaged in four or more activities to promote learning and school readiness in the past 3 days
- Percentage of children 36-59 months whose father has engaged in one or more activities to promote learning and school readiness in the past 3 days

Attendance at early childhood education:

- Percentage of children 36-59 months who are attending an early childhood education program.

Countries: Since 1995, more than 230 surveys have been conducted in more than 100 countries. MICS4 includes 65 surveys (48 national, 13 sub-national and 4 selected population) in 55 countries, 50 of which are classified as developing countries (see Appendix A).

Ages: 0-5

Administration

Location: Household

Agency: Typically by government organizations (National Statistical Offices), with technical assistance from UNICEF.

[9] "Questionnaires and Indicator List," UNICEF, accessed May 31, 2012, http://www.childinfo.org/mics4_questionnaire.html

Time: 5-10 minutes for Early Child Development subscale

Sample: Varies by country

Frequency: 1995 (MICS1), 2000 (MICS2), 2005-6 (MICS3), 2009-11 (MICS4), and every 3 years thereafter.

Data use: Global monitoring for MDGs, research

Data availability: Datasets and results for most countries are available for download through: http://www.childinfo.org/mics_available.html. Requires registration.

Instrument: Young Lives Assessment Battery

(Peabody Picture Vocabulary Test (PPVT), Child Development Assessment-Quantitative (CDA-Q); Raven's Colored Progressive Matrices (CPM); Mathematics Achievement Test (MAT); Early Grade Reading Assessment (EGRA); reading, writing, numeracy, social/emotional items developed by study)

Initiative: Young Lives: An International Study of Childhood Poverty

Type: Cross-national

Description: Young Lives is a longitudinal study of childhood poverty that includes assessments of language and cognitive development from early childhood into adulthood in order to understand the impact of poverty on children's outcomes.[10]

Learning measures:

- **Younger cohort:** Children were not given learning assessments in the first round of data collection at age 1-2. During the second round at age 4-5, they were assessed using the PPVT (language) and CDA-Q (quantitative). During the third round at age 7-8, they were assessed using the PPVT, MAT measures plus social/emotional questions and an adaptation of the Early Grade Reading Assessment (EGRA). [11]
- **Older cohort:** During the first round of data collection at age 7-8, children were assessed on intellectual capability (CPM) and three items relating to reading, writing, and numeracy. In the second round at age 11-12, they were assessed on the same three items plus the PPVT (language) and MAT (math). In the third round at age 14-15, they were assessed on the PPVT/MAT, plus a Cloze measure on reading comprehension and social/emotional questions.

Countries: Ethiopia, India (in the state of Andhra Pradesh only), Peru, Vietnam

Ages:

- Younger cohort: Surveyed at age 1-2 in 2002 (no learning assessments), age4-5 in 2006; age 7-8 in 2009
- Older Cohort: Surveyed at age 7-8 in 2002; age 11-12 in 2006; age 14-15 in 2009.

Administration

[10] Young Lives, "What we do," 2011, http://www.younglives.org.uk/what-we-do

[11] Santiago Cueto, et al, "Tracking disparities: Who gets left behind," 2011, http://www.younglives.org.uk/files/country-reports/round-3-survey-report_peru

Location: Household, with a subsample assessed at school (see note below)

Agency: Partner institutions in-country (Universities, independent research institutes, government research institutes), overseen by University of Oxford.

Time: Approximately 30-45 minutes for language/cognitive assessments

Sample: Approximately 3,000 children per country (2,000 in the younger cohort and 1,000 in the older cohort).The study uses sentinel site sampling (purposively sampled 20 high-poverty communities and then randomly sampled 150 children of the designated ages in each community). The study uses a pro-poor sample which is not nationally representative and does not include the top 10% income level. However, the study does not focus entirely on poorest children in order to understand the dynamics of poverty.

Frequency: 2002, 2006, 2009, plans for two more rounds of data collection before 2016.

Data uses: Research, national and global policy

Data availability: Available for download through the UK Economic and Social Data Service: http://www.esds.ac.uk/international/access/I33379.asp. Requires registration.

Note: In addition to these cross-country measures administered at the household level, Young Lives has conducted curriculum-related cognitive achievement tests with a sub-sample of the Younger Cohort children, as part of the school-based component that was introduced following the Round 3 household survey in 2009. In India this involved testing the Younger Cohort at age 9–10 using a single progressive test related to the curriculum coverage up to Grade 5, for three subject: Math, Telugu and English. In Vietnam, it involved testing the Younger Cohort children studying in Grade 5 at age 10–11, together with a sample of their peers, in both Math and Vietnamese. Children were tested at both the start and the end of the school year, on grade-specific, overlapping tests. In Peru, Younger Cohort children were assessed in Math and Literacy using grade-specific overlapping tests for Grades 2-6.

Instrument: Early Development Instrument (EDI)

Initiative: Various country- and local-level evaluations of school readiness.

Type: International

Description: A checklist of school readiness indicators completed by teachers in the first year of school.

Learning measures: EDI measures:

- Physical health and well-being
- Social competence
- Emotional maturity
- Language and cognitive development
- Communication skills and general knowledge

Countries: The EDI is administered in 24 countries, 14 of which are classified as developing countries (see Appendix A).

Ages: School entry (approximately age 4-6)

Administration:

Location: School

Agency: Varies by country; has been used by local/national governments, foundations, universities, UN agencies, and multilateral donors.

Time: Approximately 20-30 minutes per child

Sample: Varies by country

Frequency: Varies by country

Data use: Evaluation, research, national monitoring

Data availability: Reports from Canada cohorts are available from: http://www.offordcentre.com/readiness/reports.html. Data from developing countries are held by governments, World Bank and Aga Khan University and some are available upon request.

Instrument: Early Grade Reading Assessment (EGRA)

Initiative: Various USAID and other initiatives

Type: International

Description: A brief student assessment of literacy.

Learning measures: Foundational skills for literacy acquisition in the early grades, including letter recognition, reading simple words, understanding sentences and paragraphs, and listening with comprehension.[12]

Countries: Adapted for use in 44 countries, all of which are classified as developing countries. Plans are in progress to use EGRA in Namibia, Iraq, and Tanzania.

Ages: Grades 1-4 (Approximately age 6-10)

Administration:

Location: School

Agency: Various – RTI International, education ministries, NGOs, universities

Time: Approximately 15 minutes

Sample: Varies by project/country

Frequency: Varies by project/country

[12] "Early Grade Reading," USAID, accessed May 24, 2012, https://www.eddataglobal.org/reading/index.cfm

Data use: National or systems-level diagnostics, program evaluation, classroom-based assessment, and as a snapshot of learning in a particular school or district.

Data availability: Reports publicly available online at https://www.eddataglobal.org/.

Instrument: Early Grade Math Assessment (EGMA)

Initiatives: USAID and World Bank projects

Type: International

Description: A brief assessment of quantitative concepts.

Learning measures: Number identification, quantity discrimination, missing-number identification, word problem solving, addition and subtraction, shape recognition, and pattern extension.[13]

Countries: Adapted for use in 11 countries, all of which are classified as developing countries (See Appendix A).

Ages: Grades 1-4 (Approximately age 6-10)

Administration:

Location: School

Agency: RTI International

Time: Approximately 15 minutes

Sample: Varies by project/country

Frequency: Varies by project/country

Data use: National or systems-level diagnostics, program evaluation, research.

Data availability: Reports publicly available online at https://www.eddataglobal.org/.

Instrument: Literacy Boost Assessment

Initiative: Literacy Boost

Type: International (though primarily in South Asia and Sub-Saharan Africa)

Description: A brief formative and summative assessment designed to evaluate Save the Children's Literacy Boost program. The tool is primarily used in conjunction with teacher training and strategies to involve communities and families in early literacy. Reading skills are analyzed alongside home literacy environment information for a view of factors that influence reading in each context that informs program design, especially targeting.

[13] "Early Grade Math," USAID, accessed May 24, 2012, https://www.eddataglobal.org/math/index.cfm

Learning measurement: A simple tool that measures concepts about print, letter identification, reading simple words and texts for comprehension. Where government partners include math skills in their definition of literacy, a set of basic math competencies is also included.

Countries: Used in 15 countries: Bangladesh, Ethiopia, Guatemala, Haiti, Malawi, Mali, Nepal, Pakistan, Philippines, Mozambique, South Africa, Uganda, Vietnam, Zimbabwe, and Yemen, all of which are classified as developing countries. There are plans to expand to Afghanistan, Bhutan, Bolivia, Burundi, Indonesia, and Kenya.

Ages: Grades 1-4

Administration:

Location: School

Agency: Save the Children

Time: 15-30 minutes

Sample: Varies by project; typically a stratified random sample including a control group.

Frequency: Varies by project; often used for baseline and summative evaluation.

Data availability: Held by Save the Children (not public).

Data use: Inform instruction and community interventions, evaluate Literacy Boost program and the home/community literacy environment.

Instrument: ASER Language and Math Tools

Initiative: Annual Status of Education Report (ASER)

Type: Cross-National (India and Pakistan)

Description: ASER is a household survey that measures reading and arithmetic. ASER (meaning "impact" in Hindi and Urdu), began in India in 2005 as an evaluation tool for Pratham's literacy enrichment programs. Since then, the process and tools have been adapted for use in other countries including Pakistan. Although the name of the initiative is the same, ASER in India and ASER in Pakistan are conducted by two independent organizations.

Learning measure:

India: Basic reading and arithmetic assessments are conducted in 22 languages.[14] Reading tasks include: ability to identify letters, read simple common words, a four-line text at Grade 1 level and a longer paragraph at Grade 2 level. In India, ASER arithmetic tasks are primarily numerical. They include recognizing numbers 1 to 9, 11 to 99, ability to do basic two digit subtraction problems with borrowing and solving division problems (three digit by 1 digit). [15] The ASER measures in reading are partly based reading ability children in India are expected to have by Grade 2. The ASER arithmetic tasks are based on basic competencies that children are expected to

[14] Hindi, English, Assamese, Bengali, Gujarati, Kannada, Malayalam, Meiteilon, Marathi, Odiya, Punjabi, Tamil, Telugu, Urdu, Konkani, Boro, Nepali, Garo, Khasi, Mizo, Marra, Kok Borok.

[15] ASER Centre, "ASER: Overview, " 2011, http://www.asercentre.org/ngo-education-india.php?p=ASER+survey

know by Grade 4. In addition, to reading and arithmetic, each year there are some other domains that are also tested. The additional components/domains vary by year but have included English reading and basic vocabulary comprehension (English), "every day math" computations, general knowledge, reading maps, recognizing well known Indians etc.[16]

Pakistan: The reading tasks in Pakistan are similar to those in India.[17] The highest level of ability that is assessed is in accordance to the grade 2 level of the national curriculum. Literacy and numeracy assessments are conducted in the national language, Urdu, as well as Sindhi and Pashto in the respective regions. A literacy assessment is also administered in English. The literacy assessment includes a grade 2 level paragraph, grade 1 level sentences, and early childhood level letters and words. Numeracy is also assessed in the same manner, with the highest level assessed being division problems and the lowest level being number recognition of 1 to 9.

Countries: India and Pakistan

Ages: 5-16

Administration

Location: Household

Agency:

In India, the unit leading the ASER survey within Pratham is called ASER Centre (www.asercentre.org). The actual data collection in districts is done each year by a local organization or institution. Over 500 organizations participate each year with approximately 25,000-30,000 volunteers.

ASER in Pakistan is facilitated by Idara-e-Taleem-o-Aagahi (ITA)/South Asian Forum for Education Development (SAFED) (Pakistan). The actual survey is conducted by numerous partner local organizations and institutes, with approximately 5,000 volunteers in the 2011 survey round.

Time: Approximately 15 minutes

Sample:

India: ASER focuses on all rural districts; 30 villages in each district are randomly sampled using a rotating panel design, whereby one-third of the village rotates out and two-thirds remain the same from one year to the next. Twenty households are randomly selected in each village and all children in the age group 5 to 16 are assessed. Each year close to 300,000 household are reached in India.

Pakistan: ASER Pakistan has gradually increased in scale since its pilot in 2008 with 11 rural districts participating. There were 32 rural districts in 2010, 84 rural and 3 urban districts in 2011, and an expected 6 urban and 135 rural districts in 2012. Twenty households are surveyed in each of the 30 villages in each rural district, and not less than 27 blocks in the urban districts. The households are randomly selected in each village and block, while the villages are selected from the National Census

[16] Rukmini Banerji, Personal Communication, June 2012.

[17] "ASER: Overview," ASER Centre, accessed May 2, 2012, http://www.asercentre.org/ngo-education-india.php?p=ASER+survey

using the Population Proportionate to Size sampling technique. A total of 50,400 households were surveyed in the latest round.

Frequency: India: Annually 2006-2011, with plans to continue until there are 10 years of data; Pakistan: 2008, 2010, 2011; with plans to continue until 2015.

Data availability: India: Public reports available disaggregated by district at www.asercentre.org. The website also has a data query function. Raw data are available upon request. Pakistan: Public reports disaggregated by district available at www.aserpakistan.safedafed.org/. Raw data are available upon request.

Data use: National, provincial, and district-level policy; awareness and advocacy; academic research and evaluation.

Instrument: Uwezo Literacy and Numeracy Tools

Initiative: Uwezo

Type: Cross National

Description: Household survey in Eastern Africa that includes a measure of literacy and numeracy. Uwezo was developed based on ASER methodology.

Learning measures: Letter identification, reading simple text, number recognition, and basic arithmetic.

Countries: Kenya, Tanzania, and Uganda.

Ages: 6-16

Administration

Location: Household

Agency: Civil society organization: Hivos/Twaweza with in-country partner organizations and assistance from local volunteers and NGOs.

Time: Approximately 15 minutes

Sample: Thirty villages in each census district are sampled using a rotating panel design, whereby one-third of the village rotates out and two-thirds remain the same from one year to the next. Twenty households are randomly selected in each village, for a total of 94,800 households 158 districts in Kenya, 78,800 households in 133 districts in Tanzania, and 48,000 households in 80 districts in Uganda.

Frequency: Annually 2010 – 2013 (expected)

Data availability: Reports published online at www.uwezo.net. Cleaned data sets will be publicly available in August 2012.

Instruments: First International Comparative Study of Language, Mathematics, and Associated Factors; Second Regional Comparative and Explanatory Study

Initiative: Latin American Laboratory for Assessment (LLECE)

Type: Regional (Latin America & Caribbean)

Description: Two regional studies of learning at the primary level.[18]

Learning measurement: Mathematics, language (reading and writing) and natural science (second study and in a subset of countries only).

Countries: First study: 13 Latin American countries, all classified as developing countries. Second study: 16 countries and the Mexican State of Nuevo Leon, all classified as developing countries (See Appendix A).[19]

Ages: First study: Grades 3 and 4; Second study: Grades 3 and 6.

Administration:

Location: School

Agency: National governments

Time: 45-70 minutes for each subject (language, mathematics, and writing), administered on different days.

Sample: Two-stage (community demographics and public/private school management), weighted stratified sample**.** First study: Approximately 4,000 students per country and 54,000 total for the region. Second study: Between 8,000 – 14,000 per country, for a total of 101,000 grade 3 students and 95,000 grade 6 students.

Frequency: 1997, 2006

Data availability: Data available to research centers and institutions upon request. Reports publicly available online at www.llece.org.

Data use: National and Regional diagnostic

Instrument: PIRLS Assessment and Pre-PIRLS Assessment

Initiative: Progress in International Reading Literacy Study (PIRLS)

Type: International

Description: The International Association for the Evaluation of Educational Achievement (IEA) conducts the PIRLS, a large-scale assessment of reading comprehension processes.[20] IEA recently launched the Pre-PIRLS in 2011.[21]

[18] UNESCO LLECE, "Student achievement in Latin America and the Caribbean**:** Results of the Second Regional Comparative and Explanatory Study (SERCE)," 2008, http://unesdoc.unesco.org/images/0016/001610/161045e.pdf

[19] UNESCO LLECE, 2008 http://unesdoc.unesco.org/images/0016/001606/160660s.pdf

[20] International Association for the Evaluation of Educational Achievement, "PIRLS 2011 Assessment Framework," 2009, http://timss.bc.edu/pirls2011/downloads/PIRLS2011_Framework.pdf

Learning measurement: PIRLS measures processes of reading comprehension, such as making inferences, interpreting and integrating ideas, and evaluating content. The reading passages on the Pre-PIRLS are shorter and use simpler vocabulary and syntax than the PIRLS. The Pre-PIRLS is intended to measure reading skills in Grades 4-6, especially in countries with low literacy levels in the early grades.

Countries: PIRLS: 49 countries, 13 of which are developing countries (See Appendix A). Pre-PIRLS: 3 countries, all classified as developing countries (Botswana, Colombia, and South Africa).

Ages: PIRLS: 4th year of formal schooling; Pre-PIRLS: Grades 4-6.

Administration

Location: School

Agency: School coordinators, overseen by national research partners

Time: 80 minutes

Sample: Matrix sampling design, whereby each portion of the assessment is randomly assigned to different students in order to obtain results at the group level.

Frequency: PIRLS: 2001, 2006, 2011, plans for 2016; Pre-PIRLS: 2011, plans for 2016

Data availability: Database public and downloadable from http://timssandpirls.bc.edu/.

Data use: National-level diagnostic, international comparison, national and international policy.

Instrument: TIMSS Assessment

Initiative: Trends in International Mathematics and Science Study (TIMSS)

Type: International

Description: Assesses mathematical and science ability in the 4th and 8th years of schooling.

Learning measurement: In the 4th grade, students are tested on math concepts related to numbers, geometric shapes and measures, and data display. The science concepts tested are life science, physical science, and earth science.[22] The Grade 8 assessment includes the math concepts of numbers, algebra, geometry, and data and chance and the science concepts of biology, chemistry, physics, and earth science.[23]

Countries: 63 countries, 28 of which are classified as developing countries

Ages: Grades 4 and 8

[21] IEA, "Meeting the Needs of a Range of Countries," 2011, http://timss.bc.edu/pirls2011/prepirls.html

[22] International Association for the Evaluation of Educational Achievement, "Countries Participating in TIMSS 2011," 2011, http://timssandpirls.bc.edu/timss2011/countries.html

[23] International Association for the Evaluation of Educational Achievement, 2009.

Administration

Location: School

Agency: School coordinators, overseen by national research partners

Time: 80 minutes

Sample: Matrix sampling design, whereby each portion of the assessment is randomly assigned to different students in order to obtain results at the group level.

Frequency: 1995, 1999, 2003, 2007, 2011, plans for 2015

Data availability: Database public and downloadable from http://timssandpirls.bc.edu/.

Data use: National-level diagnostic, international comparison, national and international policy.

Instrument: PASEC Assessment

Initiative: Analysis Programme of the CONFEMEN Education Systems (PASEC)

Type: Regional (Francophone Africa)

Description: La Conférence des Ministres de l'Education des pays ayant le français en partage (CONFEMEN) has administered the PASSEC in Francophone countries since 1993.[24]

Learning measurement: PASEC assess children in reading (in French) and mathematics concepts.

Countries: The latest round includes 13 Francophone countries in Sub-Saharan Africa, all of which are classified as developing countries; (See Appendix A).

Ages: Grades 2 and 5, administered at the beginning and end of each academic year.

Administration

Location: School

Agency: National governments

Time: Varies

Sample: Stratified cluster sample, size varies by country

Frequency: Latest round, 2004-2009**:** 2004 (Chad, Guinea, Mauritania); 2005 (Benin, Cameroon, Madagascar); 2006 (Gabon, Mauritius); 2007 (Burkina Faso, Congo, Senegal); 2009 (Cote d'Ivoire, Comoros).

[24] CONFEMEN, "PASEC Comments to EFA FTI Global and Regional Activities (GRA)," 2010, http://www.educationfasttrack.org/media/GRA/GRA%20Comments%20CONFEMEN%20Learning%20outcomes.pdf

Data availability: Reports posted online at www.confemen.org. Databases available upon request at http://www.confemen.org/spip.php?article272.

Data use: National diagnostic, national/regional policy, informing education reform efforts

Instrument: SACMEQ Assessment

Initiative: Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ**)**

Type: Regional (Anglophone Africa)

Description: SACMEQ is a school-based assessment that measures proficiency in reading and mathematics in Anglophone African countries.

Learning measurement: Reading in English and mathematics.

Countries: 14 countries, all of which are classified as developing countries (See Appendix A).

Ages: Grade 6

Administration

Location: School

Agency: National governments

Time: Varies

Sample: Stratified cluster sample (a pre-determined number of schools were randomly sampled from each region). Number of respondents varies by country.

Frequency: 1995, 2000, and 2007

Data availability: Reports publicly available online at www.sacmeq.org. Database accessible upon request at http://www.sacmeq.org/data_archive.htm. Website includes data visualization feature.

Data use: Regional and national-level diagnostic, national, regional and international policy

Instrument: PISA Assessment

Initiative: Programme for International Student Assessment (PISA)

Type: International

Description: The Organisation for Economic Co-Operation and Development (OECD) conducts the PISA end-of-cycle assessment of skills essential for participation in society.

Learning measurement: Measures reading, mathematical, scientific literacy, and problem solving.

Countries: 75 countries, 32 of which are classified as developing countries.

Ages: Age 15, irrespective of grade level.

Administration

Location: School

Agency: International contractors, national governments

Time: 2 hours

Sample: Matrix sampling design

Frequency: 2000, 2003, 2006, 2009

Data availability: Database available for download at http://pisa2009.acer.edu.au/. Reports available online at www.pisa.oecd.org.

Data use: National-level diagnostic, international comparison, national and international policy.

Instrument: LAMP Assessment

Initiative: Literacy Assessment and Monitoring Programme (LAMP)

Type: International

Description: LAMP is an initiative of UNESCO Institute for Statistics (UIS) to enable countries to measure adult literacy. LAMP has been developed taking into account previous OECD experiences such as the International Adult Literacy Survey and Adult Literacy and Life Skills Survey, but the overall approach and tools have been adjusted in order to better represent the interest and needs (including the diversity of languages/scripts) of developing countries.

Learning Assessment: Measures reading and numeracy on a continuum of five levels[25]:

- Level 1: the individual has very poor skills and may, for example, be unable to determine the correct dose of medicine to give a child from the label on a package.
- Level 2: respondents can only deal with simple, clearly laid-out reading tasks. At this level, people can read but test poorly. They may have developed coping skills to meet everyday literacy demands, but they find it difficult to tackle new challenges, such as certain job skills.
- Level 3: considered a suitable minimum for coping with demands of daily life and work in a complex society. This skill level is generally required to successfully complete secondary school and enter college.
- Level 4 and 5: respondents demonstrate a good command of higher-order information processing skills.

Countries: 12 countries, all of which are classified as developing countries. Plans are in progress for the LAMP to be conducted in Nigeria and India.

[25] "LAMP – Literacy Assessment Monitoring System" UIS, accessed July 20, 2012 http://www.uis.unesco.org/literacy/Pages/lamp-literacy-assessment.aspx

Ages: 16-64

Administration

Location: Household

Agency: National governments

Time: Approximately 60-90 minutes

Sample: Sample varies by country, but is usually a probabilistic sample, representative of the target population. A stratified (often urban-rural), multi-stage (typically three or four-stage) sampling design is used: selection of enumeration areas to be sampled (typically with Probability Proportional to Size), selection of households within areas (typically systematic), and selection of individuals within households (always Simple Random Sampling among eligible members).

Frequency: Varies by country; UIS recommends cycles of 10 years"

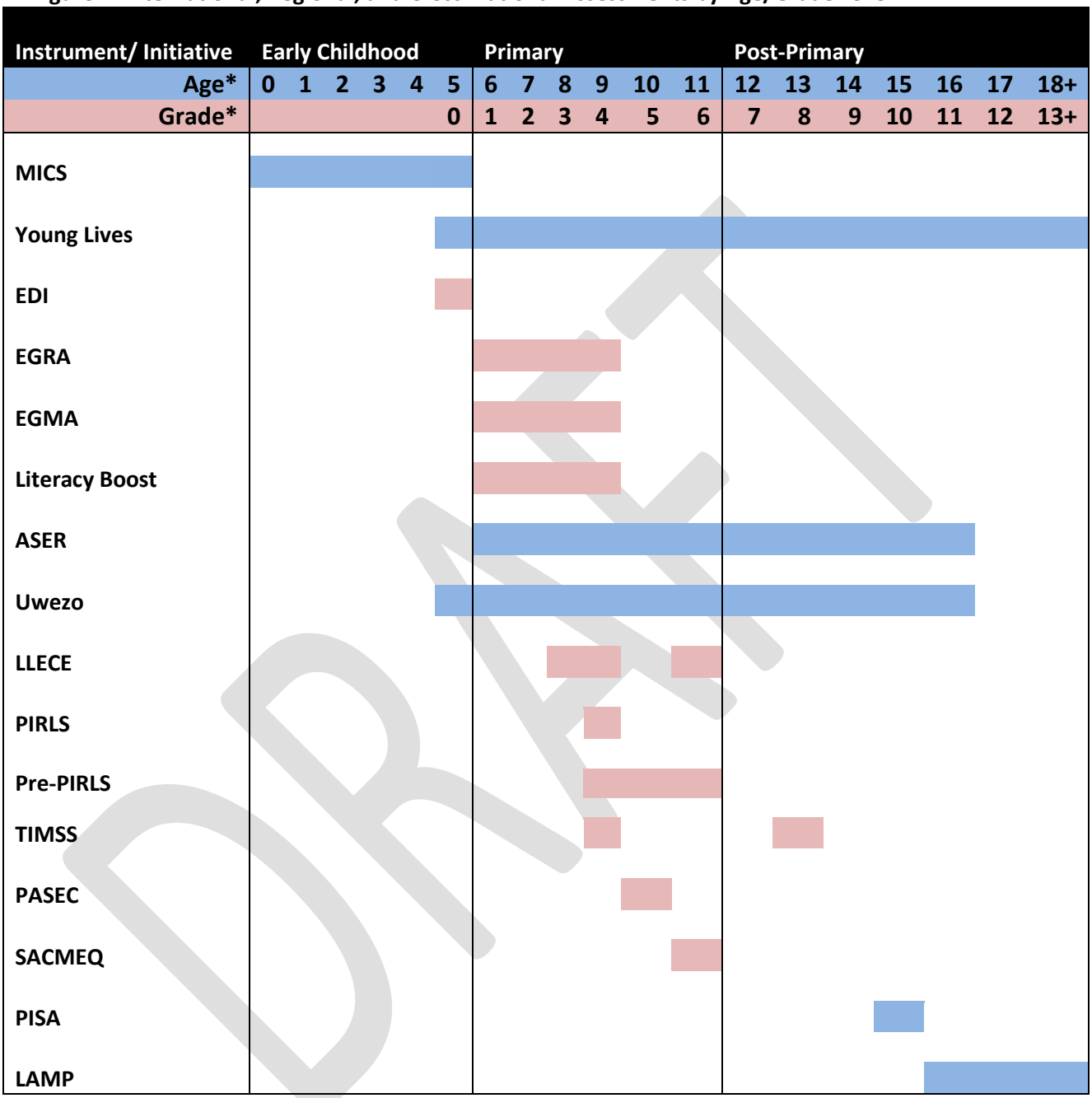
Data availability: Varies by country

Data use: National-level diagnostic, national and international policy

Characteristics of existing data

Figure 1 shows the coverage of international, regional, and cross-national assessments by age and grade level. Instruments that are administered by age area shown in blue and instruments administered by grade level are shown in red. The age and corresponding grade levels differ by country, depending on the age of school entry and number of students who start school on-time for age.

Figure 1. International, Regional, and Cross-National Assessments by Age/Grade Level



*Note: Age and grade level designations for early childhood, primary, and post-primary are context-specific. The designation here is purely for comparison of gaps and overlap.

As Figure 1 shows, the majority of assessments are conducted during the primary years. The following sections describe the strengths and gaps in existing assessments at the early childhood, primary, and post-primary levels.

Early Childhood

Internationally comparable datasets on early childhood learning is limited. There are three main sources of evidence used to directly or indirectly assess learning in early childhood:

- General statistical data on access to early childhood education programs and timely access to primary education;
- Information on opportunities for learning produced by household surveys, in particular UNICEF MICS; and
- Information produced by specialized studies and instruments such as EDI and Young Lives.

Statistical data on enrolment and timely entry to primary is part of the data regularly collected by the UIS for the population who are at least three years of age and attend an structured educational program (classified as level 0 as per the International Standard Classification of Education). The UIS data includes the following information disaggregated by gender:

- Percentage of the population who are enrolled in an early childhood education program by age (age-specific enrolment rate), or age-group (net enrolment rate).
- Percentage of the entry-age to grade 1 of primary education population who actually entered that grade (net intake rate to the first grade of primary education)
- Percentage of those entering grade 1 of primary education who previously attended an early childhood development program.

UNICEF Multiple Indicator Cluster Survey (MICS) includes questions about children's development in the domains of lanugagein the latest round. MICS has a broad geographical scope and provides some critical information on the educational situation of young children prior to school entry. Young Lives and the various initiatives assessed with the EDI are initiatives that actually measure learning but their geographical coverage is restricted.

Overall, the assessments reviewed for early childhood demonstrate a holistic approach to assessing learning. All of these initiatives/instruments include items related to multiple developmental domains, including language, numeracy, social/emotional skills, and physical development. Measurements of early childhood development are typically administered in the home, while measures of school readiness are administered upon school entry.

Primary: Early Grades and End of Cycle

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—what Dan Wagner refers to as "smaller, quicker, cheaper"[26] assessments such as ASER, EGRA, EGMA, Literacy Boost, and Uwezo—and large-scale assessments such as LLECE, PASEC, PIRLS, SACMEQ, and TIMSS. Each approach has advantages and limitations.

Simpler measures such as EGRA and ASER can capture a wide range of skills and provide usable data on the fundamentals of literacy and numeracy. The results are typically presented in language that is easily understood,

[26] Dan Wagner, *Smaller, Quicker, Cheaper: Improving Learning Assessments for Developing Countries*, Paris: UNESCO, 2011.

such as “40% of Grade 3 students cannot read a single word in English,” which catches the attention of the public and policymakers.

Because the assessments are brief, they can be adapted into different languages and scripts much more easily than a lengthy assessment. However, the adaptability makes cross-country or cross-language/script comparison unsuitable, as the instruments necessarily differ depending on the language/script of administration. For example, one indicator in the EGRA is reading fluency measured by words read per minute. In some languages where words are relatively short, a child may be able to read significantly more words per minute than in a language where words are relatively long. A simple comparison of learning across multiple languages using this indicator would favor simple orthographies over complex ones.

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. Assessments such as LLECE, PASEC, and SACMEQ are based 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 SACMEC 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. These large-scale assessments also provide a basis for international comparison, which helps governments and citizens evaluate how their education systems measure up to other countries. Because of the amount of time and resources needed to validate, administer, clean and report the data for large-scale assessments, they tend to be administered in less frequent intervals than the simpler assessments.

The simpler assessments are easier to adapt into mother tongue languages which can be a more authentic assessment of learning. For larger-scale assessments, the psychometric properties of the assessments themselves tend to be stronger, but implementation can still be weak. For example, if students are given more or less than the allotted time or allowed to ask questions of the administrator, the results can be invalidated. Large-scale assessments are more costly to administer and require more time to collect, analyze, and report data. For this reason, they tend to be administered less frequently.

In cases where each set of instruments is developed to address a particular language and script and no mechanisms are established to make the results comparable, the data cannot be aggregated at national level except for a few variables (e.g., number of children who cannot decode a single word). However, assessments that are standardized at a regional or international level leave out unique aspects of national or regional curricula and could, over time, lead to a narrowing of curriculum or “teaching to the test” among participating countries.

Post Primary

International efforts to measure post-primary learning are less developed than for primary learning. The LAMP measures literacy skills on a continuum for youth age 16 and older and is primarily administered in developing countries. The PISA measures advanced mathematical, scientific, and reading literacy and problem solving skills and is primarily administered in high-income and OECD countries. ASER and Uwezo conduct assessments with youth up to age 16, but the assessment content is based on Grade 2 standards. TIMSS is conducted in Grade 8, which is considered lower secondary by some countries and primary by others.

In order for a system of learning assessment in post-primary to be useful, it must encompass more than just literacy but be flexible enough to include the variety of realities that youth face during the transition from adolescence to adulthood (e.g. parenthood, work, higher education). Administering the TIMSS or PISA in countries where basic literacy and numeracy is low creates a floor effect, whereby little useful data are produced because so few students get any of the questions correct. Similarly, assessments such as ASER and Uwezo could have a ceiling effect when administered to 16-year-old youth, but because they are administered to both in- and out-of-school youth, they do provide some useful data on literacy in the entire population. While the LAMP produces much-needed data on reading and numeracy level, this assessment alone is not sufficient to determine whether youth and adults have the knowledge and skills to compete in a modern economy. For post-primary learning, a more robust system of measuring learning is needed beyond what is currently available.

Recently, the OECD has developed a new measure of adult learning, the Programme for the International Assessment of Adult Competencies (PIAAC). This assessment is based on past surveys such as the National Adult Literacy Survey (US) and the International Adult Literacy Survey and the Adult Literacy and Life Skills Survey. The PIAAC is a household-based survey of youth and adults age 16-64 and will measure literacy and numeracy skills and the ability to solve problems in technology-rich environments. Currently, there are plans to administer the PIAAC in OECD countries and the Russian Federation only.

Building on Current Momentum

There is clearly momentum for measuring learning globally. National governments, civil society organizations, multilateral organizations, NGOs, and donors have demonstrated a willingness to fund and conduct large-scale assessments of learning.

There are several initiatives in process to address these gaps and build on current efforts. The World Bank and UNESCO are each working to increase national government capacity to measure learning.

- The Student Assessment domain of World Bank's System Assessment and Benchmarking for Education Results (SABER) works to help developing countries strengthen their systems for student assessment. SABER is an evidence-based program that provides tools and resources to help countries systematically examine and strengthen the performance of their education systems.[27]
- The UNESCO General Education Quality Diagnostic was developed to systematically assess the education systems in UNESCO member countries and identify barriers to quality teaching and learning.

There are also efforts to track learning outcomes on a global scale. Some countries have set national education goals based on international education benchmarks. For example, the Organisation of Ibero-American States in its *Metas Educativas 2021* sets goals for the percentage of students in the top and bottom tiers of international assessments such as PIRLS, TIMSS, and PISA.[28] The UNESCO Institute for Statistics, UNESCO International Institute for Educational Planning (IIEP), and the Center for Universal Education at the Brookings Institution have launched three initiatives to examine ways to align and synthesize existing efforts.

[27] World Bank, "SABER – Student Assessment" 2011, http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/0,,contentMDK:22808581~menuPK:282391~pagePK:148956~piPK:216618~theSitePK:282386,00.html

[28] Organisation de Estados Iberoamericanos, "Metas Educativas 2021," 2011, http://www.oei.es/metas2021/libro.htm

- The UNESCO Institute for Statistics (UIS) has launched the *Observatory of Learning Outcomes* to track learning achievement and measurement worldwide.[29] 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 Center for Universal Education at the Brookings Institution and UNESCO, through its statistical branch (UIS) have convened a *Learning Metrics Task Force* (LMTF) facilitate a shared vision for learning and discuss the feasibility of global learning metrics. The task force will engage a wide range of global actors to make recommendations for assessment of learning at the early childhood, primary, and post-primary levels.

- 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.[30] 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.

- UNESCO is working with an interagency team to develop the Holistic Early Childhood Development Index (HECDI) to monitor global progress toward EFA Goal 1.[31] The final version of this tool will contain child outcome indicators that may include early development and/or learning.

Conclusion

Much progress has been made in increasing access to school and measuring learning globally. However, several gaps exist that make it impossible to obtain an accurate estimation of learning worldwide.

The majority of initiatives to measure learning at the primary level are conducted in schools, thereby leaving out children who are of school age but not enrolled in school. This methodology provides some information on how much children are learning in school, but on a global scale could actually decrease the incentive for improving access to education if it means aggregate learning levels in schools will decrease. Filmer, Hassan, and Pritchett[32] propose a model for a "Millennium Learning Goal" in which access and learning are both factored into achievement of the goal. The authors first estimated the "learning trajectory" in each country based on incremental increases in standardized test scores across grade levels. Based on the proportion of a given age cohort that reaches the learning goal threshold (through test performance) and the proportion of the age cohort

[29] UNESCO, "Observatory of Learning Outcomes: Design Summary," 2011, http://www.uis.unesco.org/Education/Pages/observatory-of-learning-outcomes.aspx

[30] UNESCO IIEP, "Sharing knowledge on quality of education in Africa," 2011, http://www.iiep.unesco.org/en/news/single-view.html?tx_ttnews%5Btt_news%5D=778

[31] UNESCO, "Better Monitoring in Sight for Young Children," 2011, http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/education-for-all/single-view/news/better_monitoring_in_sight_for_the_well_being_of_young_children/

[32] Deon Filmer, Amer Hassan & Lant Pritchett, "A Millennium Learning Goal: Measuring Real Progress in Education," 2006, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=982968

that finishes a given grade, the authors estimate national learning levels in 11 countries and provide a model for global measurement toward learning goals.

Current assessment efforts leave out the most marginalized. While significant efforts have been made to collect learning assessment data globally, these efforts are disproportionally in high- and middle-income countries. Furthermore, some developing countries participating in global assessments only submit their higher-performing districts for assessment. For example, the PISA is administered in only three districts in China (Hong Kong, Macao, and Shanghai) and two in India (Tamil Nadu and Himachal Pradesh). While the rankings for Shanghai and Hong Kong are extremely high (Shanghai is first and Hong Kong is fourth internationally in reading), these figures are not representative of the country as a whole. Twenty-five developing countries do not participate in any multi-country assessments, potentially due to the cost of participating in these assessments. Most of the non-participating countries are also among the lowest in the world in terms of gross domestic product (GDP).

These gaps in current assessment systems make it impossible to measure the current state of learning worldwide. Each assessment captures part of the global learning picture, but the result is a fragmented snapshot and not a true picture of learning globally.

Bibliography

"ASER: Overview," ASER Centre, 2011. http://www.asercentre.org/ngo-education-india.php?p=ASER+survey

Banerji, Rukmini. Personal Communication, June 2012.

Burnett, Nicholas. "A learning goal for education? Scoping for MDG2 post-2015": Report to Hewlett Foundation/Wellspring Advisors, Washington, DC: Results for Development Institute, 2012.

CONFEMEN. "PASEC Comments to EFA FTI Global and Regional Activities (GRA)." 2010. http://www.educationfasttrack.org/media/GRA/GRA%20Comments%20CONFEMEN%20Learning%20outcomes.pdf

Cuerto, Santiago, Javier Escobal, Mary Penny and Patricia Ames. "Tracking disparities: Who gets left behind." 2011. http://www.younglives.org.uk/files/country-reports/round-3-survey-report_peru

Filmer, Deon, Amer Hassan and Lant Pritchett. "A Millennium Learning Goal: Measuring Real Progress In Education." 2006. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=982968

Guadalupe, Cesar. "Learning outcomes: the key challenge." Montreal: UIS, 2012

International Association for the Evaluation of Educational Achievement. "PIRLS 2011 Assessment Framework." 2009. http://timss.bc.edu/pirls2011/downloads/PIRLS2011_Framework.pdf

International Association for the Evaluation of Educational Achievement. "Countries Participating in TIMSS 2011." 2011. http://timssandpirls.bc.edu/timss2011/countries.html

International Association for the Evaluation of Educational Achievement. "Meeting the Needs of a Range of Countries." 2011. http://timss.bc.edu/pirls2011/prepirls.html

"LAMP – Literacy Assessment Monitoring System" UIS, accessed July 20, 2012 http://www.uis.unesco.org/literacy/Pages/lamp-literacy-assessment.aspx.

"Literacy Boost." Save the Children. 2011, accessed http://www.savethechildren.org/site/c.8rKLIXMGIpI4E/b.7084483/

Organisation de Estados Iberoamericanos. "Metas Educativas 2021." 2011. http://www.oei.es/metas2021/libro.htm

UNESCO, "Better Monitoring in Sight for Young Children," 2011, http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/education-for-all/single-view/news/better_monitoring_in_sight_for_the_well_being_of_young_children/

UNESCO. *EFA Global Monitoring Report.* Paris: UNESCO, 2011.

UNESCO. “Observatory of Learning Outcomes: Design Summary.” 2011. http://www.uis.unesco.org/Education/Pages/observatory-of-learning-outcomes.aspx

UNESCO IIEP. “Sharing knowledge on quality of education in Africa.” 2011. http://www.iiep.unesco.org/en/news/single-view.html?tx_ttnews%5Btt_news%5D=778

UNESCO LLECE. “Student achievement in Latin America and the Caribbean”: Results of the Second Regional Comparative and Explanatory Study (SERCE). 2008. http://unesdoc.unesco.org/images/0016/001610/161045e.pdf

UNICEF. MICS Optional Questions and Question Modules. 2008. http://www.childinfo.org/mics3_questionnaire.html

USAID. “Early Grade Reading.” 2011. https://www.eddataglobal.org/reading/index.cfm

USAID. “Early Grade Math.” 2011. https://www.eddataglobal.org/math/index.cfm

Wagner, Dan “Smaller, Quicker, Cheaper: Improving Learning Assessments for Developing Countries.” 2011. Paris: UNESCO.

Young Lives. “What we do.” 2011. http://www.younglives.org.uk/what-we-do

Appendix A: International, Regional, and Cross-National Assessments in Developing Countries

	Country[33]	MICS4	Young Lives	EDI	EGRA	EGMA	Literacy Boost	ASER	Uwezo	LLECE	PIRLS	Pre-PIRLS	TIMSS	PASEC	SACMEQ	PISA	LAMP
East Asia and Pacific																	
	American Samoa																
	Cambodia			1	1												
	China										1		1			1	
	Fiji																
	Indonesia	1		1							1		1			1	
	Kiribati																
	Korea, Dem. Rep.																
	Lao PDR	1															1
	Malaysia												1			1	
	Marshall Islands																
	Micronesia, Fed. Sts																
	Mongolia	1															1
	Myanmar																
	Palau																
	Papua New Guinea				1												
	Philippines			1	1		1										
	Samoa																
	Solomon Islands																
	Thailand	1											1			1	
	Timor-Leste				1												
	Tonga				1												
	Tuvalu																
	Vanuatu				1												
	Vietnam	1	1	1	1		1										1
East and Central Asia																	

[33]Note: This table provides data on all countries in which an assessment is administered; in some cases, this may be at a sub-national level or for purposes other than national diagnosis.

Appendix A: International, Regional, and Cross-National Assessments in Developing Countries

	Country[33]	MICS4	Young Lives	EDI	EGRA	EGMA	Literacy Boost	ASER	Uwezo	LLECE	PIRLS	Pre-PIRLS	TIMSS	PASEC	SACMEQ	PISA	LAMP
	Albania															1	
	Armenia												1				
	Azerbaijan										1		1			1	
	Belarus	1															
	Bosnia and Herzegovina	1															
	Bulgaria										1					1	
	Georgia										1		1			1	
	Kazakhstan	1											1			1	
	Kosovo			1													
	Kyrgyz Republic															1	
	Latvia															1	
	Lithuania										1		1			1	
	Macedonia, FYR	1											1				
	Moldova	1		1												1	
	Montenegro	1														1	
	Romania										1		1			1	
	Russian Federation										1		1			1	
	Serbia	1											1			1	
	Tajikistan																
	Turkey												1			1	
	Turkmenistan	1															
	Ukraine	1											1				
	Uzbekistan	1															
Latin America & Caribbean																	
	Antigua and Barbuda																
	Argentina	1			1					1						1	
	Belize	1															
	Bolivia									1							

Appendix A: International, Regional, and Cross-National Assessments in Developing Countries

	Country[33]	MICS4	Young Lives	EDI	EGRA	EGMA	Literacy Boost	ASER	Uwezo	LLECE	PIRLS	Pre-PIRLS	TIMSS	PASEC	SACMEQ	PISA	LAMP
	Brazil			1	1					1						1	
	Chile			1						1			1			1	
	Colombia									1	1	1				1	
	Costa Rica	1								1						1	
	Cuba	1								1							
	Dominica																
	Dominican Republic									1							
	Ecuador									1							
	El Salvador									1							1
	Grenada																
	Guatemala				1		1										
	Guyana				1												
	Haiti				1		1										
	Honduras				1					1	1		1				
	Jamaica	1		1	1												1
	Mexico			1						1						1	
	Nicaragua				1	1											
	Panama	1								1						1	
	Paraguay									1							1
	Peru		1	1	1					1						1	
	St. Kitts and Nevis																
	St. Lucia	1															
	St. Vincent/Grenadines																
	Suriname	1															
	Uruguay	1														1	
	Venezuela, RB									1						1	
Middle East & North Africa																	
	Algeria	1															

Appendix A: International, Regional, and Cross-National Assessments in Developing Countries

	Country[33]	MICS4	Young Lives	EDI	EGRA	EGMA	Literacy Boost	ASER	Uwezo	LLECE	PIRLS	Pre-PIRLS	TIMSS	PASEC	SACMEQ	PISA	LAMP
	Djibouti																
	Egypt, Arab Rep.				1												
	Iran, Islamic Rep.										1		1				
	Iraq	1				1											
	Jordan			1		1							1			1	1
	Lebanon	1											1				
	Libya																
	Morocco				1	1					1		1				1
	Occupied Palestinian Territory	1											1				1
	Syrian Arab Republic												1				
	Tunisia	1											1			1	
	Yemen, Rep.				1		1						1				
South Asia																	
	Afghanistan	1			1												1
	Bangladesh				1		1										
	Bhutan	1															
	India		1		1			1								1	
	Maldives																
	Nepal	1			1		1										
	Pakistan	1		1	1		1	1									
	Sri Lanka																
Sub-Saharan Africa																	
	Angola				1												
	Benin													1			
	Botswana										1	1	1		1		
	Burkina Faso													1			

Appendix A: International, Regional, and Cross-National Assessments in Developing Countries

Country[33]	MICS4	Young Lives	EDI	EGRA	EGMA	Literacy Boost	ASER	Uwezo	LLECE	PIRLS	Pre-PIRLS	TIMSS	PASEC	SACMEQ	PISA	LAMP
Burundi				1									1			
Cameroon													1			
Cape Verde																
Central African Republic	1															
Chad	1												1			
Comoros													1			
Congo, Dem. Rep.	1			1	1											
Congo, Rep													1			
Côte d'Ivoire													1			
Eritrea																
Ethiopia		1		1		1										
Gabon													1			
Gambia, The	1			1												
Ghana	1			1								1				
Guinea																
Guinea-Bissau	1															
Kenya	1			1	1			1						1		
Lesotho														1		
Liberia				1	1											
Madagascar	1			1									1			
Malawi				1	1	1								1		
Mali	1			1	1	1										
Mauritania	1												1			
Mauritius													1	1	1	
Mayotte																
Mozambique			1	1		1								1		
Namibia														1		1
Niger				1												1

Appendix A: International, Regional, and Cross-National Assessments in Developing Countries

Country[33]	MICS4	Young Lives	EDI	EGRA	EGMA	Literacy Boost	ASER	Uwezo	LLECE	PIRLS	Pre-PIRLS	TIMSS	PASEC	SACMEQ	PISA	LAMP
Nigeria	1			1												
Rwanda				1	1											
São Tomé and Principe																
Senegal				1									1			
Seychelles														1		
Sierra Leone	1			1												
Somalia	1															
South Africa				1		1					1	1		1		
South Sudan	1															
Sudan	1															
Swaziland	1													1		
Tanzania								1						1		
Togo	1															
Uganda				1		1		1						1		
Zambia				1	1									1		
Zimbabwe						1								1		
Total Countries Participating	**50**	**4**	**14**	**44**	**11**	**15**	**2**	**3**	**16**	**13**	**3**	**28**	**13**	**14**	**32**	**12**

MICS4: Early Child Development Index (ECDI)

EC1.	1. HOW MANY CHILDREN'S BOOKS OR PICTURE BOOKS DO YOU HAVE FOR (*name*)?	None ..... 00 Number of children's books ..... 0 __ Ten or more books 10	
EC2.	2. I AM INTERESTED IN LEARNING ABOUT THE THINGS THAT (*name*) PLAYS WITH WHEN HE/SHE IS AT HOME. DOES HE/SHE PLAY WITH [A] HOMEMADE TOYS (SUCH AS DOLLS, CARS, OR OTHER TOYS MADE AT HOME)? [B] TOYS FROM A SHOP OR MANUFACTURED TOYS? [C] HOUSEHOLD OBJECTS (SUCH AS BOWLS OR POTS) OR OBJECTS FOUND OUTSIDE (SUCH AS STICKS, ROCKS, ANIMAL SHELLS OR LEAVES)? *If the respondent says "YES" to the categories above, then probe to learn specifically what the child plays with to ascertain the response*	Y N DK Homemade toys..... 1 2 8 Toys from a shop ..... 1 2 8 Household objects or outside objects 1 2 8	

Source: MICS4, 2009

Young Lives

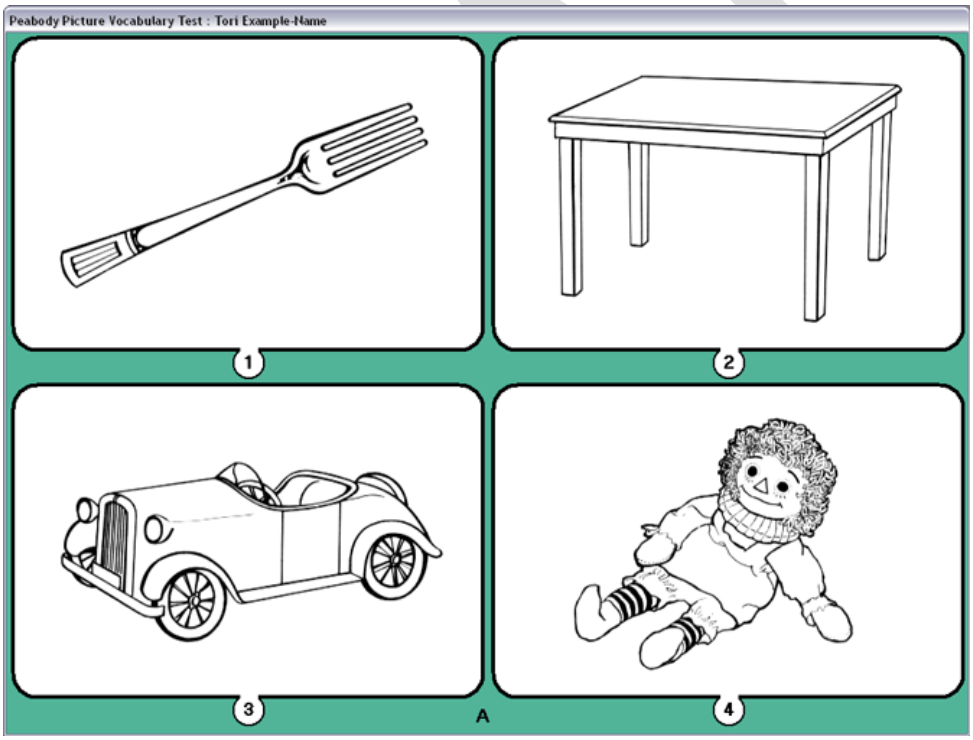
Reading, writing, and numeracy items

- The reading item required children to read three letters ('T, A, H'), one word ('hat'), and one sentence ('the sun is hot') in India, Ethiopia, and Vietnam. In Peru, the reading item required the children to read the letters 'N, A, P'; the word '*pan*' (Spanish for 'bread'); and the sentence '*El pan es rico*' (Spanish for 'the bread is tasty').
- The writing item asked the children to write a simple sentence ('I like dogs') which was spoken out loud by the examiner.
- The numeracy item required children to solve a basic multiplication problem (2 x 4).

Source: Cueto et. al., 2009

Peabody Picture Vocabulary Test (PPVT)

Child is asked "point to the table," etc.



Source: PPVT-R

Early Development Instrument (EDI)

Section B - Language and Cognitive Skills

How would you rate this child's:	very good/ good	average	poor/ very poor	don't know
1. ability to use language effectively in English	○	○	○	○
2. ability to listen in English	○	○	○	○
3. ability to tell a story	○	○	○	○
4. ability to take part in imaginative play	○	○	○	○
5. ability to communicate own needs in a way understandable to adults and peers	○	○	○	○
6. ability to understand on first try what is being said to him/her	○	○	○	○
7. ability to articulate clearly, without sound substitutions	○	○	○	○

Source: EDI, 2010

Early Grade Reading Assessment (EGRA)

(From administrator's guide: Child is presented with a page with letters only).

Sample Assessment Design: Letter Name Knowledge

Show the child the sheet of letters in the student stimuli booklet. Say:

Here is a page full of letters of the alphabet. Please tell me the NAMES of as many letters as you can-- not the SOUNDS of the letters, but the names.

For example, the name of this letter *[point to A]* **is "A"**

Let's practise: tell me the name of this letter *[point to V]:*
If the child responds correctly say: **Good, the name of this letter is "VEE."**
If the child does not respond correctly, say: **The name of this letter is "VEE."**
Now try another one: tell me the name of this letter *[point to L]:*
If the child responds correctly say: **Good, the name of this letter is "ELL."**
If the child does not respond correctly, say: **The name of this letter is "ELL."**

Do you understand what you are to do?

When I say "Begin," please name the letters as quickly and carefully as you can. Start here and continue this way. *[Point to the first letter on the row after the example and draw your finger across the first line].* **If you come to a letter you do not know, I will tell it to you. Otherwise I will keep quiet & listen to you. Ready? Begin.**

Start the timer when the child reads the first letter. Follow along with your pencil and ***clearly*** *mark any incorrect letters with a slash (/). Count self-corrections as correct. If you've already marked the self-corrected letter as incorrect, circle the letter and go on.* ***Stay quiet,*** *except when providing answers as follows: if the child hesitates for 3 seconds, provide the name of the letter, point to the next letter and say "****Please go on****." Mark the letter you provide to the child as incorrect. If the student gives you the letter sound, rather than the name, provide the letter name and say: ["****Please tell me the NAME of the letter****"]. This prompt may be given only once during the exercise.*

AFTER 60 SECONDS SAY, "stop." Mark the final letter read with a bracket ( ]).

Early stop rule: *If the child does not give a single correct response on the first line, say "****Thank you!****", discontinue this exercise, check the box at the bottom, and go on to the next exercise.*

Example : A v L

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	
L	i	h	R	S	y	E	O	n	T	*(10)*
i	e	T	D	A	t	a	d	e	w	*(20)*
h	O	e	m	U	r	L	G	R	u	*(30)*
g	R	B	E	i	f	m	t	s	r	*(40)*
S	T	C	N	p	A	F	c	a	E	*(50)*
y	s	Q	A	M	C	O	t	n	P	*(60)*
e	A	e	s	O	F	h	u	A	t	*(70)*
R	q	H	b	S	i	g	m	i	L	*(80)*
L	i	N	O	e	o	E	r	p	X	*(90)*
N	A	c	D	d	I	O	j	e	n	*(100)*

Time remaining on stopwatch at completion (number of SECONDS) : ☐

Check this box if the exercise was discontinued because the child had no correct answers in the first line. ☐

Appendix B: Sample Items for International, Regional, and Cross-National Assessments

Source: EGRA Administrator's Guide, 2011

Appendix B: Sample Items for International, Regional, and Cross-National Assessments

Early Grade Math Assessment (EGMA)

(From administrator's guide. Child is presented with a page with numbers only).

Task 1: Number Identification | Sheet A | 60 seconds

Here are some numbers. I want you to point to each number and tell me what the number is. I am going to use this stopwatch and will tell you when to begin and when to stop.
- *[point to first number]* Start here. Are you ready? . . . Start.
- What number is this ?

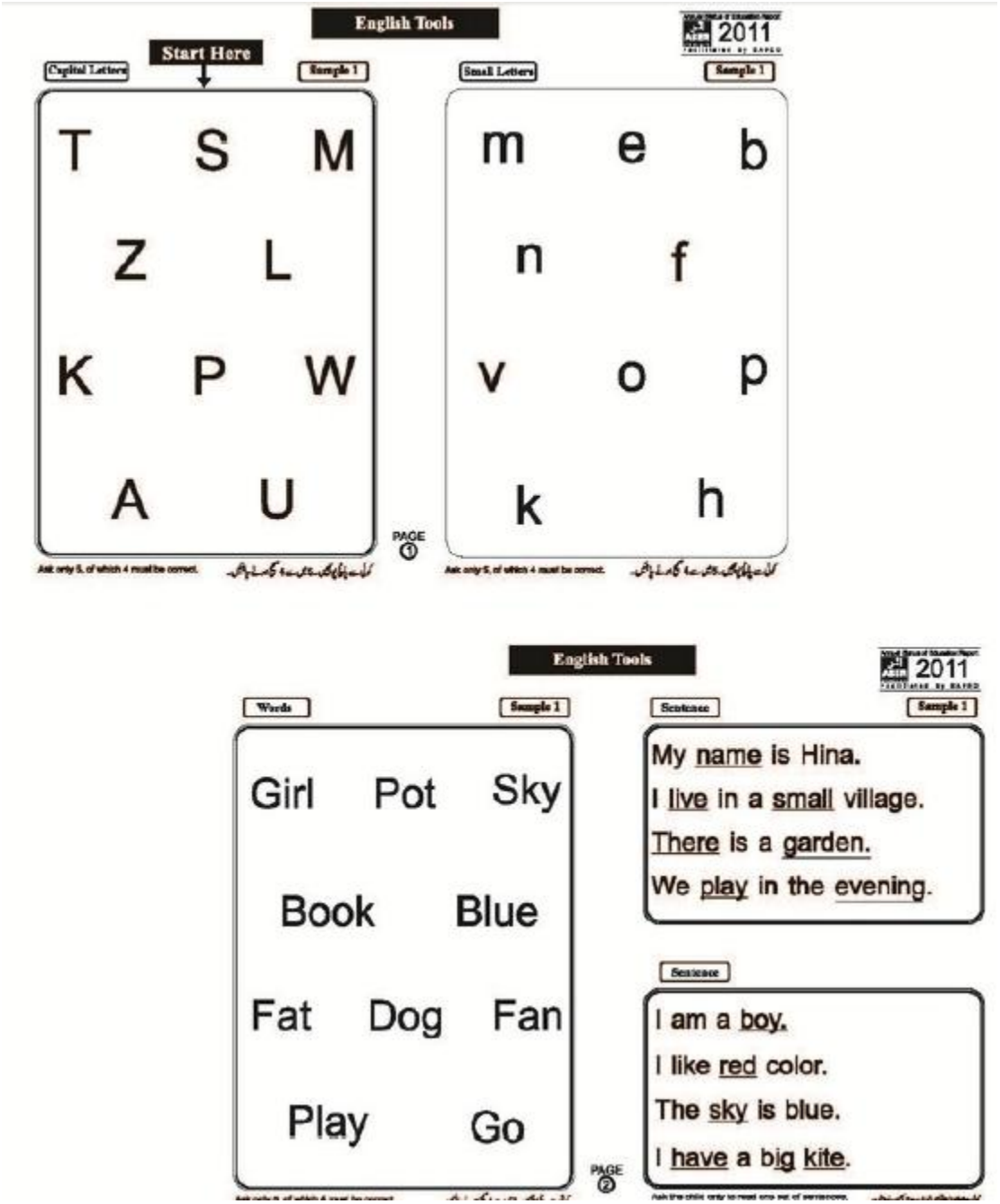
(/) Incorrect or no response
( ]) After the last number read

- If the time on the stopwatch runs out (60 seconds).
- If a child stops on a number for 5 SECONDS.

					Tot. Cum.
4	10	28	58	807	(5)
94	368	30	106	17	(10)
9	39	14	711	83	(15)
423	34	72	245	77	(20)
187	52	22	19	33	(25)
646	12	64	49	301	(30)

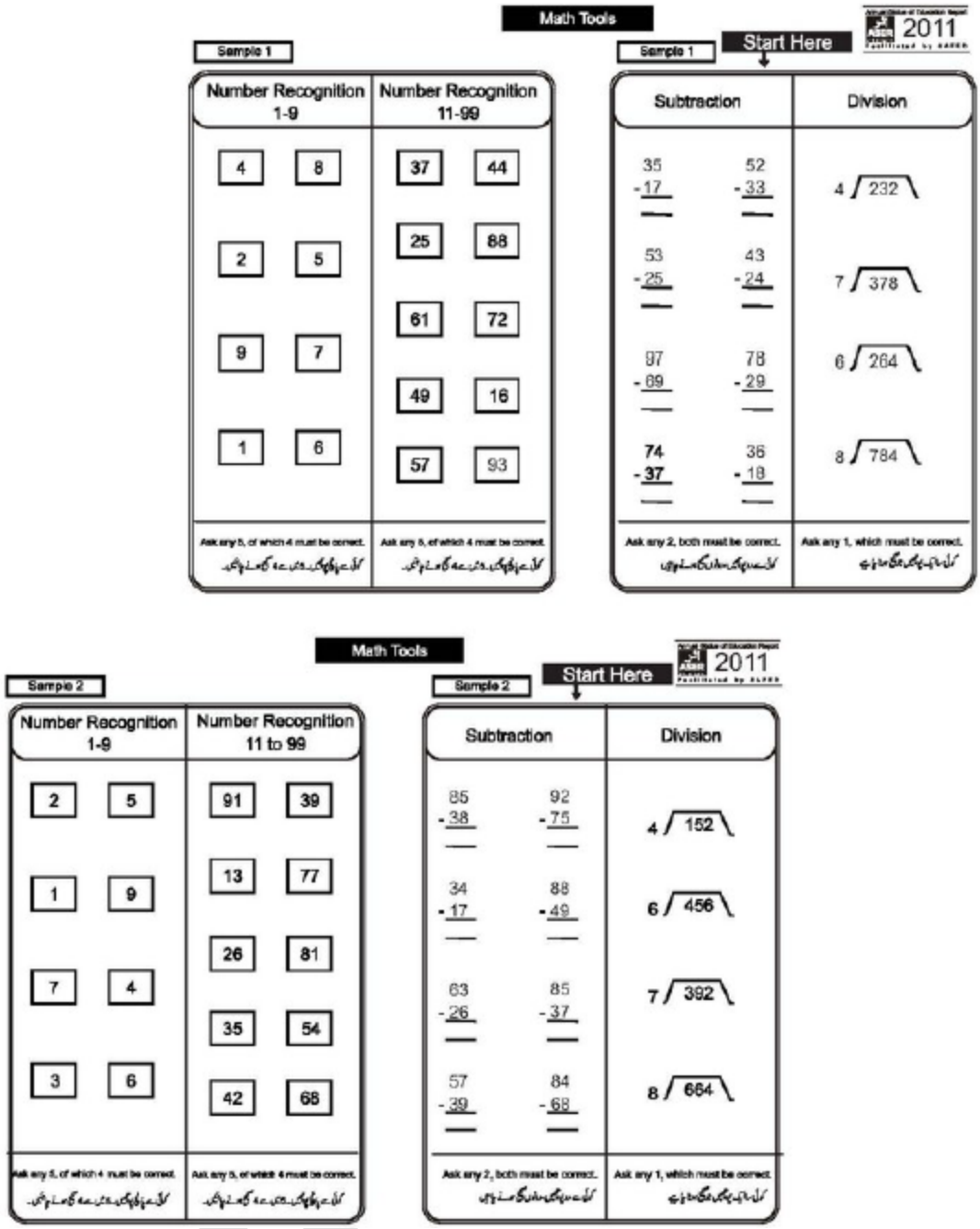
Source: EGMA Administrator's Guide, 2012

ASER Language and Math Tools



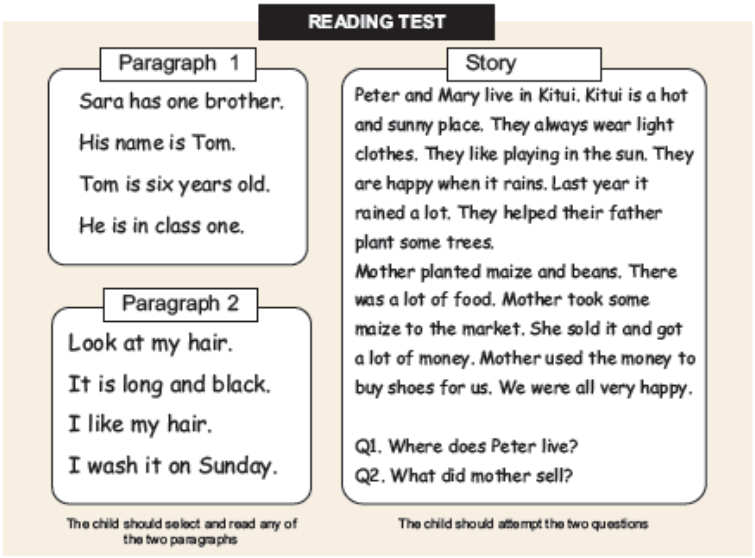
Source: ASER Pakistan English Literacy Tool, 2011

Appendix B: Sample Items for International, Regional, and Cross-National Assessments



Source: ASER Pakistan Math Tool, 2011

Uwezo Literacy Tools



Source: Uwezo Kenya English Reading Test, 2011

Appendix B: Sample Items for International, Regional, and Cross-National Assessments

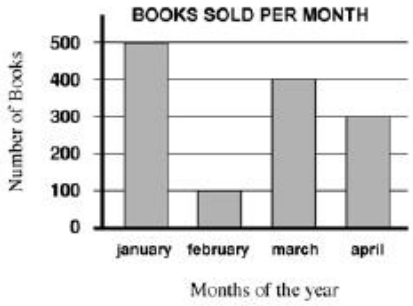
Second Regional Comparative and Explanatory Study, Latin American Laboratory for Assessment (LLECE)

Level I - 3rd Grade Mathematics

Example 1. Books sold per month

19

The following graph shows the number of books sold in a bookstore during the first



Which month shows the highest book sale?months of the year:

- A) January.
- B) February.
- C) March.
- D) April.

Source: SERCE, 2007

Level II - 6th Grade Reading

Example 3. The perfect horse

THE BEST HORSE

Once upon a time there was a prince and a princess who were getting married. All they needed was a horse to pull the wedding chariot. So they decided to convoke all the kingdom's horses to select the one that would be most suitable.

The first one to arrive was a very fast horse called Chocolate. However, when he was tested the chariot toppled over. Then came a horse called Pampered. His proud owner said "This one is not too fast, but he is terribly smart: he understands everything you tell him". Then they shouted 'Giddy-up, giddy-up', but the horse did not move. Several other horses came and went. The prince and the princess were desperate.

Suddenly a farmer showed up leading this tiny horse. "This little horse is called Mini" he said, "and he is very fast, strong and smart". Everybody burst out laughing. It was a pony! To everyone's surprise, however, he turned out to be the perfect horse. The day of the wedding he was crowned as the best horse in the world.

Authoress: Alejandra Rintá, a Fourth Grade student.

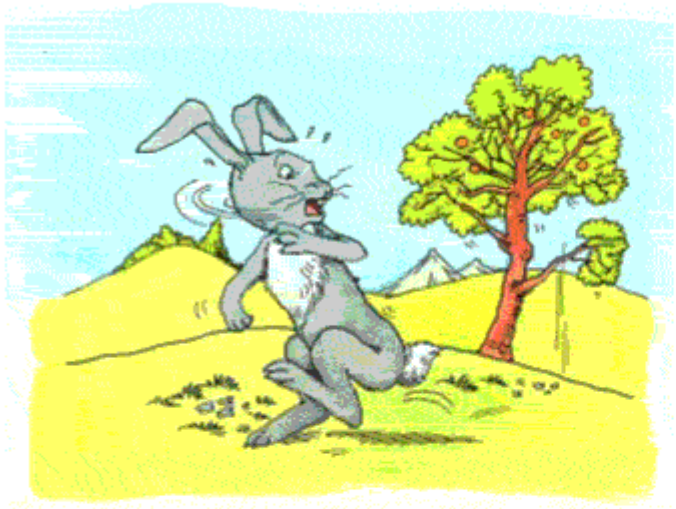
16 According to his owner, the second horse was

- A proud.
- B smart.
- C small.
- D fast.

Source: SERCE, 2007

PIRLS Assessment and Pre-PIRLS Assessment

Rabbit Raises the Earthquake Alarm



Question #1
Question #2
Question #3
Question #4
Question #5

by: Rosalind Kerven

There was once a rabbit who was always worrying. "Oh dear," he muttered all day long, "oh deary, deary me."

His greatest worry was that there might be an earthquake. "For if there was," he said to himself, "whatever would become of me?"

He was feeling particularly anxious about this one morning, when suddenly an enormous fruit fell down from a nearby tree—*CRASH!*—making the whole earth shake.

The rabbit leaped up.

"Earthquake!" he cried.

Literary Subscale Passage and Items

Rabbit Raises the Earthquake Alarm -

Question:

What made the whole earth shake?

A. an earthquake
B. an enormous fruit
C. the fleeing hares
D. a falling tree

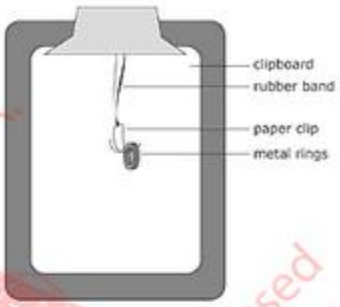
Answer:

The correct answer is B.

Source: PIRLS 2001

Trends in International Mathematics and Science Study (TIMSS) – 8th Grade Science

The figure below shows a clip on a clipboard holding a rubber band. A pape clip is bent to form a hook for hanging metal rings on the lower end of the rubber band.



Using the above set up, 10 metal rings, and a ruler, you will find out how the length of the rubber band changes as more and more rings are hung on it. Hang the metal rings onto the paper clip one by one and measure the length of the rubber band for each new ring.

A. Write your measurements in the table below. Remember to write the heading for each column.

B. Graph your results below.

C. When there are 2 rings on the paper clip and 3 more are then added, how much longer does the rubber band become?

The rubber band becomes ____ cm longer.

D. Describe how the rubber band changed in length as more and more rings were added.

E. What do you think would be the length of the rubber band if you could add 2 more rings than you have been given?

I think the total length of the rubber band might be _____ cm.

F. Why do you think this would happen?

Source: TIMSS 2003

Appendix B: Sample Items for International, Regional, and Cross-National Assessments

Analysis Programme of the CONFEMEN Education Systems (PASEC)

Needed

Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)

Needed

Programme for International Student Assessment (PISA)

Figure 1 shows changing levels of Lake Chad, in Saharan North Africa. Lake Chad disappeared completely in about 20,000 BC, during the last Ice Age. In about 11,000 BC it reappeared. Today, its level is about the same as it was in AD 1000.

Figure 2 shows Saharan rock art (ancient drawings or paintings found on the walls of caves) and changing patterns of wildlife

Source: Past Worlds: The Times Atlas of Archaeology, Times Books Limited 1988

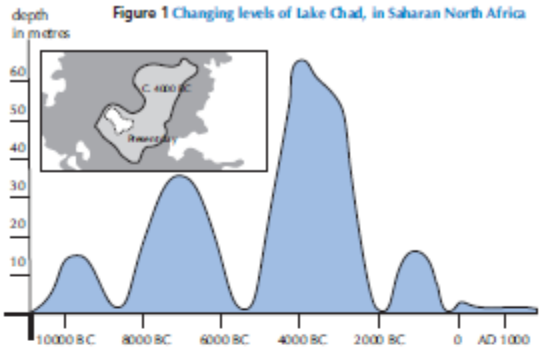


Figure 1 Changing levels of Lake Chad, in Saharan North Africa

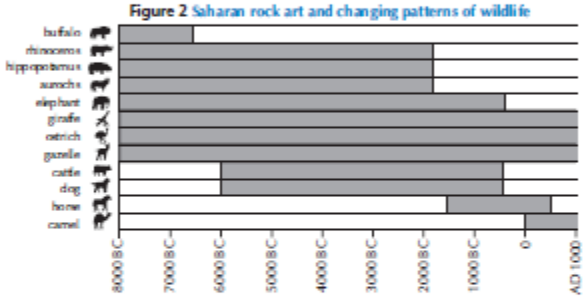


Figure 2 Saharan rock art and changing patterns of wildlife

Use the above information about Lake Chad to answer the questions below.

QUESTION 1.1

What is the depth of Lake Chad today?

A. About two metres.

B. About fifteen metres.

C. About fifty metres.

D. It has disappeared completely.

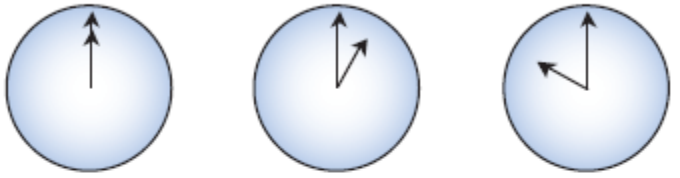
E. The information is not provided.

Source: PISA, 2006

MATHEMATICS UNIT 11: INTERNET RELAY CHAT

Mark (from Sydney, Australia) and Hans (from Berlin, Germany) often communicate with each other using "chat" on the Internet. They have to log on to the Internet at the same time to be able to chat.

To find a suitable time to chat, Mark looked up a chart of world times and found the following:



Greenwich 12 Midnight | Berlin 1:00 AM | Sydney 10:00 AM

QUESTION 11.1

At 7:00 PM in Sydney, what time is it in Berlin?

Answer: .....

QUESTION 11.2

Mark and Hans are not able to chat between 9:00 AM and 4:30 PM their local time, as they have to go to school. Also, from 11:00 PM till 7:00 AM their local time they won't be able to chat because they will be sleeping.

When would be a good time for Mark and Hans to chat? Write the local times in the table.

Place	Time
Sydney	
Berlin	

Source: Pisa, 2006

Literacy Assessment and Monitoring Programme (LAMP)

MEDCO ASPIRIN 500

INDICATIONS: Headaches, muscle pains, rheumatic pains, toothaches, earaches. RELIEVES COMMON COLD SYMPTOMS.

DOSAGE: ORAL. 1 or 2 tablets every 6 hours, preferably accompanied by food, for not longer than 7 days. Store in a cool, dry place.

CAUTION: Do not use for gastritis or peptic ulcer. Do not use if taking anticoagulant drugs. Do not use for serious liver illness or bronchial asthma. If taken in large doses and for an extended period, may cause harm to kidneys. Before using this medication for chicken pox or influenza in children, consult with a doctor about Reyes Syndrome, a rare but serious illness. During lactation and pregnancy, consult with a doctor before using this product, especially in the last trimester of pregnancy. If symptoms persist, or in case of an accidental overdose, consult a doctor. Keep out of reach of children.

INGREDIENTS: Each tablet contains
500 mg acetylsalicicylic acid.
Excipient c.b.p. 1 tablet.
Reg. No. 88246

0 67736 11079

Made in Canada by STERLING PRODUCTS, INC.
1900 Industrial Blvd., Montreal, Quebec H5J 3P1

Reprinted by permission

What is the maximum number of days you should take this medicine?

In total, how many bottles are in the two full cases?

Appendix B: Sample Items for International, Regional, and Cross-National Assessments

Source: UNESCO
http://www.uis.unesco.org/StatisticalCapacityBuilding/Workshop%20Documents/Education%20workshop%20dox/2011%20Kingston/4_UIS_Info_session_on_LAMP.pdf