

LITERACY DATA: MORE, BUT NOT ALWAYS BETTER

Since 2000, when reducing illiteracy was included as a specific indicator in Education for All's Dakar Framework for Action, there has been substantial interest in accurately measuring literacy rates around the world. This interest has intensified recently as the international development community looks forward to the post-2015 global education agenda, with its growing emphasis on measuring learning in key domains, including literacy.

In response, over the past decade, international civil society organizations have developed a variety of new externally administered literacy assessments, which are designed to gauge literacy skills among different ages and groups of individuals. In this brief, we discuss the myriad sources of data about literacy and what they mean for consumers of literacy data.

UNESCO Institute for Statistics (UIS) compiles adult literacy data reported by national education agencies; it is generally assumed that literacy rates gathered directly from countries are comparable. In fact, there is wide diversity in the way that countries define literacy, which means that historical data on literacy rates are likely not comparable cross-nationally.

Recognizing these limitations, cross-national literacy data increasingly come from external assessments. Although some of these assessments are specifically designed to be standardized cross-nationally, many are not.

Literacy is a fundamental human right and the foundation for lifelong learning

-UNESCO

A major concern is that literacy data from various external assessments cannot be reliably compared to those from other assessments due to numerous differences in their design, including:

- Different sample populations – differences in age, region, language, and in- or out-of-school status
- Varying levels of text difficulty and (in some cases) the challenges of standardizing texts across languages
- Varying definitions of literacy and ability levels, which are not aligned to one another

Additionally, external literacy assessments are rarely designed to assess students' writing skills, or their literacy in multiple languages, meaning they likely do not capture students' full range of literacy abilities. Ultimately, we argue that consumers of literacy data must take all literacy data with a healthy dose of skepticism—the specific assessment design and definition of literacy used significantly influences the conclusions we can draw about the state of literacy in a given context.

PART I – UNDERSTANDING LITERACY DATA

Literacy Definitions around the World

No single definition of literacy exists, and definitions of what it means to be literate have evolved over time. In 1958, UNESCO adopted a resolution that defined literacy as “the ability to both read and write, with understanding, a simple statement related to one’s everyday life” (UNESCO 1958). Since then, scholars and practitioners have elaborated this definition to involve a broader sense of competencies including the ability to make sense of (and critique) the world, interpret symbolic systems, and communicate with others.

Despite this broadening perspective, most international development frameworks continue to adopt a narrower, skill-based, definition of literacy. UNESCO’s 2006 Global Monitoring Report (GMR) defines literacy as: “a set of tangible skills – particularly the cognitive skills of reading and writing – that are independent of the context in which they are acquired and the background of the person who acquires them” (UNESCO EFA GMR 2006, p. 149).

However, many countries adopt more specific definitions of literacy for national purposes, which are hardly standardized. For example, in some countries, literacy is defined as “the ability to read and write simple sentences,” while in others, such as Angola, Bosnia and Herzegovina, Kenya, and Sudan it is defined as “the ability to read easily or with difficulty a letter or a newspaper” (UNESCO 2006, p. 157).

Most surprising are countries that make no reference to actual abilities to read and write – instead, literacy is defined entirely in reference

to participation in the formal education system. For example, in Malaysia, individuals are defined as literate if they are above age 10 and have been to school in any language. In Mali, an individual is classified as illiterate if he or she never attended school – regardless of whether he or she can read or write. The distinction is important, as 2013 Beekungo results find that less than half of second grade students in most regions of Mali are able to pass a basic reading exam (EPDC 2014).

In fact, some nations’ definitions of literacy are language- and even residence- specific – such as in China, where the dominant language is non-alphabetic and literacy is defined as knowledge of 2,000 characters in urban areas and 1,500 characters in rural areas (UNESCO 2006, p. 157). These varying definitions, which serve as the basis of UNESCO Institute of Statistics data on literacy, make understanding a country’s literacy rate difficult, and cross-country comparisons practically impossible.

Literacy Data: From Self-Reports to External Assessment

There are many sources of data on literacy; we focus on those that are specifically designed to produce generalized knowledge about the state of literacy in a given country or context. The purpose of these literacy data is to understand the level and type of literacy in a sample or population of individuals, often in order to make comparisons with other nations, assess educational quality, or inform future policy decisions.

The earliest cross-national data on literacy came from self-reports – since the 1960s, UNESCO Institute of Statistics has compiled data from censuses and statistical bureaus. In the vast majority of countries, individuals are simply asked “Can you read and write?” – often with no reference to type or difficulty of content. As UNESCO (2014) explains: “In most countries, there are no other measures – just a simple count based on information gathered in a household survey or census. Answer “yes” and join the ranks of the so-called literates. Answer “no” and you are considered illiterate” (UNESCO 2014). Case studies suggest that self-reporting tends to result in inflated – sometimes dramatically so – literacy rates, as respondents often overestimate or misrepresent their reading and writing abilities (Schaffner 2005).

Additionally, self-reports are often completed by heads-of-household, who may not actually be familiar with individual family members’ literacy levels. Moreover, self-reports provided by the UNESCO Institute of Statistics shed little insight into the current state of literacy in a given nation. Most self-reports cover all adults aged 15 and older which means they tend to reflect the nation’s prior educational infrastructure, rather than the current level of access or quality.

In response to the many shortcomings of self-reported literacy data, the international community has increasingly moved away from self-reports to externally validated and administered assessments. These external assessments seem to offer many advantages. Perhaps most importantly, rather than defining literacy as a binary status (i.e. literate/illiterate),

they can assess a spectrum of abilities, which ultimately offers a more nuanced classification of individuals’ literacy levels.

External assessments can also test students’ abilities in various literacy sub-skills. For example, some students may not be able to recognize letters, while others may be able to pronounce words, but not comprehend them. Literacy assessments now exist to test individuals’ abilities on foundational literacy skills, including: letter-sound knowledge, phonemic awareness and reading comprehension, which will likely lead to better teaching and learning in the long run.

These external assessments would at first suggest a more objective and externally validated indicator of literacy; however, in practice they have many limitations. First, although the ability to write is included in most definitions of literacy, in practice, most literacy assessments exclude writing. This focus on reading makes sense, as tests of reading and comprehension are much easier to standardize. However, it also means that the discussion about literacy rates is even narrower than our definition of literacy. Moreover, as we discuss below, external assessments are also extremely susceptible to the definition of literacy used, the specific skills tested, and the population sampled.

Sources of External Literacy Assessment Data

External literacy assessments can be classified into three major types: cross-national, standardized exams administered in schools

(both regional and international); community-based assessments administered in homes; and, nation-specific assessments typically administered in schools. In addition, we examine data from the Early Grade Reading Assessments (EGRA), which measures students' early reading skills such as word recognition and phonemic awareness.

Community-based Household Assessments

UWEZO, ASER and Beekungo are the three major community-implemented household assessments that assess literacy. They test basic skills, including letter name knowledge, simple word reading, sentence reading, and basic comprehension. Because they assess students in their homes, they are able to assess literacy skills of children who are both in and out of school, which is something that no other assessment does. They also cover a broad age range, testing children aged 5-16, rather than those in specific school grades.

Nonetheless, the levels of literacy tested are quite low – the highest level of reading tested in any of the exams aligns to the ability to answer two questions on a two to three sentence story related to daily life. This low benchmark emerges from the exams' focus on assessing functional literacy; however, it is likely too low to be a meaningful indicator of reading in an academic setting, which is an important factor in students' ultimate educational attainment. Additionally, although UWEZO efforts are coordinated across multiple East African nations, literacy assessments are pegged to the 2nd grade curriculum in their national context. This means that the results of these exams are not strictly comparable to those in other nations, given the differences in national curricula.

Regional and International Standardized Assessments

One of the major advantages of regional and international student assessments, which distinguishes them from other assessments, is that they are standardized cross-nationally, thereby allowing for meaningful comparisons of student skills and competencies. The major regional exams with available literacy data are the: Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), the Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC) and the Segundo Estudio Regional Comparativo y Explicativo (SERCE). Regional exams such as these offer standardization across many countries in the same language (English, French and Spanish, respectively).

Despite the obvious advantage of comparison between countries participating in regional studies, there are also limitations. These assessments only test students who are in school, and exam content is pegged to a certain grade. For example, SACMEQ does not provide insight into the literacy abilities of children who dropped out of school before 6th grade, which means that its indicators significantly overestimate literacy in the population as a whole.

At the international level, a growing number of countries also participate in the Progress for International Reading and Literacy Study (PIRLS), an internationally-standardized student assessment conducted at 4th grade, which benchmarks students' literacy to four levels.

However, very few low- and middle-income nations participate in PIRLS, simply because even the lowest literacy benchmark in PIRLS is above the literacy abilities of the vast majority of their students. The lowest benchmark for fourth graders in PIRLS is the ability to “recognize, locate, and reproduce explicitly stated details from the texts, particularly if the details were close to the beginning of the text,” as well as making straightforward inferences. The overwhelming majority of students in many low-income countries would fall into this level; as a result, PIRLS offers little insight into to students’ spectrum of literacy abilities or the specific needs of students in many low- and middle-income nations.

Nation-Specific Assessments

In addition to standardized regional and international assessments, there are two other widely cited types of nation-specific data on literacy. The first is national exams – many countries engage in standardized exams of samples of their students with the goal of assessing students’ educational achievement and examining their progress over time. Because of their close linkages to national curriculum, these exams are not intended to be comparable cross-nationally. The data from such exams is also hard to access, often only provided in national reports that do not give much detail into the definitions of literacy used or the specific test items. As a result, making meaning of reported statistics can be quite difficult. Also, literacy levels may be linked narrowly to the national curriculum, which can mean distinguishing generalized literacy abilities as distinct from mastery of content can be difficult.

Finally, the Early Grade Reading Assessment (EGRA) is a rapid-diagnostic tool widely adopted in the international development community. It takes between 15 and 20 minutes per student and is administered individually and orally. EGRA is a fast and relatively cost-effective method for assessing students’ literacy sub-skills, including phonemic awareness, letter name knowledge, and fluency, as well as basic comprehension.

However, EGRA makes it clear that comparisons should not be made across languages or countries, which limits its ability to produce generalized data on the state of literacy in a country. Additionally, one of the most commonly cited indicators produced by EGRA is the *average number of words read per minute*. Rather than viewing literacy as a relatively static indicator of ability (i.e., can or cannot read) EGRA’s indicators emphasize the rate of reading as a more meaningful definition of literacy (i.e., the speed at which one can read). Although this ultimately may be a more meaningful indicator of overall literacy abilities, its findings are difficult to compare to other exams, where there is little documentation concerning how long students are given to read passages and complete comprehension questions. In fact, until now, the only meaningful indicator of literacy that has been considered comparable across countries and contexts is a zero score vs. non-zero score.

PART II – KEY CHALLENGES IN UNDERSTANDING LITERACY DATA

The growing number of cross-national assessments, with their varying definitions of literacy, mean that consumers of literacy data

face real difficulties when seeking answers to even basic questions about literacy, such as: what percent of children at a given age can read?

TABLE 1. OVERVIEW OF LITERACY DATA

Data Type	Sources	Countries	Years	Site	Ages	Language	Literacy Levels
Self-Report	UNESCO UIS	Worldwide	1970-Present	Household	Ages 15+	Not asked	Self-reports of literate or not
Household-based	UWEZO	Kenya, Tanzania, Uganda	2009-2012	Household	Ages 6-16	English, Kiswahili (Uganda – only English)	Oral administration: 5 levels. Highest level is students who can read and comprehend a Grade 2 level text (answer 2 questions)
	ASER	India, Pakistan	2005-Present	Household	Ages 5-16	Local language and English	5 levels for English; 6 levels for local language. Highest level in local languages is students who can read a Grade 2 level text; for English, highest level is students who can read sentences
	Beekungo	Mali	2012	Household	Ages 6-14	Local language and French	6 levels. Highest level is children who can correctly answer 2 questions about the story
Regionally and Internationally Standardized	SACMEQ	14 countries in East Africa	1995, 2000, 2007	School-based	G6	English	8 levels ranging from pre-reading to analytical or inferential reading
	SERCE	16 developing nations in Latin America and Caribbean	2006	School-based	G3 G6	Spanish	4 levels (Composite score of mean 500) related to students' ability to locate information in text, interpret, and paraphrase

TABLE 1. OVERVIEW OF LITERACY DATA (CONTINUED)

Data Type	Sources	Countries	Years	Site	Ages	Language	Literacy Levels
	PIRLS	49 countries, 13 developing nations	2001, 2006, 2011	School-based	G4	National language	4 internationally standardized benchmarks, low to advanced. Emphasis on acquiring information and evaluating language
Rapid Diagnostic Tools	EGRA	11 countries	2006-Present	School-based	G1 G3	19 languages including English, Arabic, French and Spanish	Oral assessment: sub-tests include: letter name knowledge; phonemic knowledge; oral reading fluency with comprehension; listening comprehension; dictation. Reading passage is high 1st grade – low 2nd grade level.

A major problem is simply that each assessment samples from a distinct population – each is targeted to children of different ages, either students in school or all children in and out of school, tested in different languages, and sometimes, in only some regions of the country. For example, EGRA data is not always available nationwide, as it tends to be implemented in conjunction with donor-funded projects. UWEZO assesses students both in and out of school, while PIRLS and SACMEQ only test students in school. This means that we cannot compare indicators of literacy rates across these assessments because they are drawn from different sample populations. As a result, it is very difficult to compare or validate data across sources.

Another issue is that each assessment uses slightly different definitions of literacy – for example, each assessment may focus on only a set of literacy sub-skills, or base its definition of reading competency on substantially easier texts than another assessment. Moreover, it is important to note that even while most definitions of literacy include writing, few external assessments actually assess or report writing abilities.

Each assessment also classifies students into varying levels of literacy (four, five, six, up to eight levels of literacy) and it is simply not clear how a given level of literacy on one assessment maps onto levels from another assessment.

This is made more complicated by the fact that in some tests, a literacy level maps onto a particular skill (e.g., ability to identify letters, ability to read words) while in other tests, students are given a composite score (e.g., 500) based on multiple test items, and this score refers to generalized abilities.

Moreover, even within the same assessment, results may not be comparable across samples. For example, EGRA assessments cannot be compared across languages even within the same country. Most EGRA tests are targeted to a high 1st grade or low 2nd grade level, while in reality, word difficulty is very context-specific (e.g. students may be familiar with British English terms, rather than American English). Until now, no scientific process exists for standardizing EGRA assessments across languages, and country contexts. This means that literacy data from different assessments cannot be reliably compared.

Linguistic diversity also poses distinct challenges for garnering comparable indicators of literacy cross-nationally. Ideally, literacy assessments should be able to capture students' linguistic abilities in multiple languages to best assess their literacy levels. However, regional and international literacy assessments rarely test more than one language, and when they do, their findings are not always consistent across languages either because students' abilities differ across languages or tests are not designed similarly across languages.

Finally, despite the growing number of assessments, availability of data remains an issue. Although substantial data on literacy exists for some countries, there are still major gaps in coverage of literacy rates for some countries,

populations, and sub-groups. This means that UNESCO self-reported literacy data is still the only indicator of literacy rates available in many nations, despite the move towards external assessments in the international community.

Similarly, even when comparisons across tests are possible because they assess similar samples of students (similar ages, grades, languages, same country), getting comparable data is not easy. Compiling findings from different exams requires mining through various report documents, posted on different websites and downloading and filtering Excel data. In response, EPDC's Learning Outcomes Data has standardized and compiled findings from various reports and sources to make data on learning outcomes comparable as possible.

Ultimately, we argue that we must take all literacy data with a healthy dose of skepticism – the specific definition of literacy used in assessments influences the conclusions we draw about the state of literacy in a given context. Moreover, given the variety of sources of data on literacy, obtaining reliable and consistent indicators of literacy rates worldwide is still extremely difficult. In the next section, we examine how difficult it is to get consistent data on literacy by examining the case of Uganda.

PART III – A CASE STUDY OF UGANDA

This section highlights the difficulties we face in trying to assess the state of literacy in a specific country context. We examine literacy in Uganda because it is one of the few countries where data on literacy is available from multiple sources for roughly the same student population in the same year.

Literacy data is available for 3rd and 6th grade students drawn from UWEZO, EGRA, SACMEQ and Uganda's national assessment, the National Assessment of Progress in Education (NAPE). EGRA data is only available for the Central region of the country, and can be compared to UWEZO data, which disaggregates data by subnational region. All tests assess literacy in English, and data on literacy in local languages is also available in EGRA and NAPE. Table 1 shows the various sources of data available:

TABLE 2. AVAILABLE UGANDA ASSESSMENTS

Test	Grades	Language	Years
UWEZO	G3, G6	English	2011
EGRA	G2, G3	English, Luganda, Lango	2009
SACMEQ	G6	English	2007
NAPE (Uganda)	G3, G6	English; 9 local languages	2010, 2012



We start with a basic question that educators, international development observers, and government professionals may all be interested in: how well can Ugandan children read? To begin to answer this question with data, we specify our age range and select one specific literacy skill: what percent of Ugandan 3rd graders can read connected text (i.e., words and sentences)?

As Table 3 on the next page shows, there are at least three sources of data on this question –all designed to answer this exact question. Both the 2010 UWEZO and 2010 NAPE have indicators on students' abilities to read words and read sentences. EGRA data is only available for the Central region, and is therefore compared to findings from UWEZO in the Central region. Table 3 shows side-by-side comparisons for Grade 3 data and Table 4 shows data for Grade 6.

TABLE 3. SIDE BY SIDE COMPARISONS – UGANDA GRADE 3 LITERACY DATA

Test	EGRA	UWEZO	UWEZO	NAPE
UWEZO	2009	2010	2010	2010
EGRA	G3	G3	G3	G3
SACMEQ	Central	Central	Nationwide	Nationwide
NAPE (Uganda)	Both	Both	Both	Both
EGRA	English	English	English	English
EGRA	76.6% (1+ WPM)	63.8% (Words) 22.1% (Sentences)	45.8% (Words) 12.0% (Sentences)	45.0% (Complete words) 50.4% (Read and complete sentences)

TABLE 4. SIDE BY SIDE COMPARISONS – UGANDA GRADE 6 LITERACY DATA

Test	UWEZO	SACMEQ	NAPE
Year	2010	2007	2010
Grade Level	G6	G6	G6
Region	Nationwide	Nationwide	Nationwide
Gender	Both	Both	Both
Language	English	English	English
Measure of Literacy	95.2% (Words) 82.3% (Sentences)	79.7% (Basic Reading: Match phrases and complete sentences) 54.2% (Reading for Meaning: Interpret sentence and paragraph level texts, read on and read back to make meaning)	53.4% (Reading and describing activities in a picture) 31% (Reading a story)

Different Definitions of Literacy

The tables above shed light onto how difficult it is to understand the state of literacy even in a clearly identified population. In Table 3, among 3rd graders in the Central region, EGRA finds that 76.6% of students can read at least one word, while UWEZO finds that only 63.8%

of children of this age can read words. This discrepancy may be due to the fact that UWEZO is based in households, and so likely includes students who never made it to Grade 3.

Nationwide, Table 2 compares students from UWEZO to the Ugandan NAPE. The findings are

quite different from one another, in part because the NAPE’s literacy indicators are not clearly defined. One of the sub-competencies tested is the ability to “complete words” and a second is the ability to “read and complete sentences”. While only 45% of students can complete words, 50.4% can read and complete sentences. In most tests, the ability to read sentences would be expected to be lower than words, so the NAPE’s findings are confusing. It is possible that the sentences used in the NAPE exam are closely linked to the curriculum, and that students may be able to succeed on these questions due to memorization rather than general literacy skills. Alternatively, these competences could be based on texts of very different difficulty, but the difficulty of the texts is never explained.

Similarly, as Table 4 shows, in the 6th grade NAPE exam, 53.4% were able to “read and describe activities in a picture,” while only 31% were adequate or advanced at “reading a story.” The differences in these two activities is never made clear in any report or public documentation.

The discrepancies found between tests may be due to varying definitions of literacy used: some tests emphasize skills and competencies associated with the practice of reading, while others focus on comprehension. SACMEQ’s definition of “Reading for Meaning” is defined as “reads on or reads back in order to link and interpret information located in various parts of a text.” In contrast, NAPE’s definition is based on a composite of indicators and incorporates skills such as matching and comprehension. The take away is simply that how we define literacy matters in terms of our understanding of who is literate in a given context.

Language Matters

The Uganda case also provides an opportunity to examine literacy within the same population of students in more than one language. Table 5 extracts indicators on literacy in three languages, from two sources: NAPE and EGRA. The table examines different definitions of literacy in two separate regions — the Luganda-speaking region and the Lango-speaking region within the same test. We do not suggest that the indicators from the two tests be directly compared in this case –

TABLE 5. SIDE BY SIDE COMPARISON – LITERACY IN DIFFERENT LANGUAGES (GRADE 3)

Source and Indicator	Luganda Region ¹		Lango Region ²	
	Luganda	English	Lango	English
NAPE: Literacy Proficiency Rate	60.7	74.9	25.6	30.3
EGRA: Can Read 1+ Word of Connected Text	74.2	76.6	35.7	49.0
EGRA: Can Read Text and Answer 1+ Comprehension Question	62.5	8.6	34.7	1.2

¹ NAPE Includes districts: Buikwe, Kayunga, Mpigi, Mukono, Wakiso

² NAPE Includes districts: Amolatar, Apac, Dokolo, Lira, Otuke, Oyam

as they could be drawing from Luganda speaking students in the Central region or the Northern region. However, we can make comparisons across languages within the same test.

The EGRA data show that among the same students, phonemic awareness is higher in English than local languages, but comprehension is substantially lower. Additionally, the literacy assessment used in Uganda's NAPE finds that students are generally more literate in English than their local languages, excepting a few regions and languages, which suggests that some regions are better at teaching literacy in local languages than others – but also makes it difficult to draw conclusions about the state of literacy generally.

Definitions of literacy depend crucially on what skills are tested and what language they are tested in. It is important to test a variety of skills, including comprehension, and a variety of languages, as EGRA Uganda report notes: “children able to “read” in English do not understand the meaning of the words they have learned” (Piper 2010, p. 5).

What is Needed?

The more we measure literacy, the more data we have – but unless definitions and levels of literacy are both clarified and standardized, the

more confusion these numerous indicators can create. Although it is unreasonable to expect the international educational community to come to complete consensus on any one definition of literacy, researchers should map out how different assessments' literacy levels map onto others, and which sub-set of literacy skills each assessment is testing.

On the other hand, it is also possible that the move to measuring literacy along a spectrum of skills offers room for advancing understandings of literacy – instead of forcing literacy assessments to map onto older definitions of “literate/illiterate,” it may be time to re-conceptualize how we report on literacy statistics. For example, countries may increasingly seek to report the percent of students who are literate at either a basic level or an advanced level, to give more clarity and nuance to their own literacy statistics.

Additionally, despite our focus on generalized and comparable statistics, we recognize that generalized statistics on literacy do little to help teachers in the classroom. Our focus on internationally comparable statistics should not detract from an emphasis on developing locally responsive for literacy assessments and school-based interventions.

REFERENCES

- EPDC. 2014. EPDC Database, Learning Outcomes Data. Available at: <http://www.epdc.org/>
- Piper, Benjamin. 2010. "Uganda Early Grade Reading Assessment Findings Report: Literacy Acquisition and Mother Tongue." Kampala, Uganda: Makerere University Institute for Social Research.
- Shaffner, Julie. 2005. UNESCO. "Measuring literacy in developing country household surveys: issues and evidence." Background paper prepared for the Education for All Global Monitoring Report 2006 Literacy for Life. Available at: <http://unesdoc.unesco.org/images/0014/001462/146285e.pdf>.
- UNESCO. 1958. Recommendations Concerning the International Standardization of Educational Statistics. Paris, UNESCO.
- UNESCO 2006. "Literacy for Life." Education for All Global Monitoring Report. Paris, France: UNESCO.
- UNESCO. 2014. "LAMP – Literacy Assessment and Mapping Programme." Available at: <http://www.uis.unesco.org/literacy/Pages/lamp-literacy-assessment.aspx>