

UNDERSTANDING WHAT WORKS IN ORAL READING ASSESSMENTS

Recommendations from donors, implementers and practitioners

UNDERSTANDING WHAT WORKS IN ORAL READING ASSESSMENTS

Recommendations from donors, implementers and practitioners



UNESCO

The constitution of the United Nations Educational, Scientific and Cultural Organization (UNESCO) was adopted by 20 countries at the London Conference in November 1945 and entered into effect on 4 November 1946. The Organization currently has 195 Member States and 10 Associate Members.

The main objective of UNESCO is to contribute to peace and security in the world by promoting collaboration among nations through education, science, culture and communication in order to foster universal respect for justice, the rule of law, and the human rights and fundamental freedoms that are affirmed for the peoples of the world, without distinction of race, sex, language or religion, by the Charter of the United Nations.

To fulfil its mandate, UNESCO performs five principal functions: 1) prospective studies on education, science, culture and communication for tomorrow's world; 2) the advancement, transfer and sharing of knowledge through research, training and teaching activities; 3) standard-setting actions for the preparation and adoption of internal instruments and statutory recommendations; 4) expertise through technical co-operation to Member States for their development policies and projects; and 5) the exchange of specialised information.

UNESCO is headquartered in Paris, France.

UNESCO Institute for Statistics

The UNESCO Institute for Statistics (UIS) is the statistical office of UNESCO and is the UN depository for global statistics in the fields of education, science and technology, culture and communication.

The UIS was established in 1999. It was created to improve UNESCO's statistical programme and to develop and deliver the timely, accurate and policy-relevant statistics needed in today's increasingly complex and rapidly changing social, political and economic environments.

The UIS is based in Montreal, Canada.

Published in 2016 by:

UNESCO Institute for Statistics
P.O. Box 6128, Succursale Centre-Ville
Montreal, Quebec H3C 3J7
Canada
Tel: (1 514) 343-6880
Email: uis.publications@unesco.org
<http://www.uis.unesco.org>
©UNESCO-UIS 2016

ISBN 978-92-9189-187-0

Ref: UIS/2016/LO/TD/4

DOI: <http://dx.doi.org/10.15220/978-92-9189-187-0-en>

This publication is available in Open Access under the Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) license (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository (<http://www.unesco.org/open-access/terms-use-ccbysa-en>).

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities or concerning the delimitation of its frontiers or boundaries.

The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO and do not commit the Organization.

FRONT AND BACK COVER PHOTOS: © Dana Schmidt/The William and Flora Hewlett Foundation, except front top left
© Hannah-May Wilson, PAL Network

The UNESCO Institute for Statistics (UIS) led a collaborative project to formulate recommendations to guide practitioners when selecting, conducting and using oral reading assessments. The aim is to highlight basic principles that should be applied in the different stages of oral reading assessments—from planning and design to implementation and use of the resulting data. The recommendations are drawn from a collection of articles, which can be found online in the ebook, *Understanding What Works in Oral Reading Assessments*, at <http://www.uis.unesco.org>

Suggested citation

UNESCO Institute for Statistics (UIS) (2016). *Understanding What Works in Oral Reading Assessments: Recommendations from Donors, Implementers and Practitioners*. Montreal: UNESCO Institute for Statistics.

Support for this initiative was generously provided by the Global Partnership for Education and the William and Flora Hewlett Foundation.

Contributors

Organization	Author
Australian Council for Educational Research (ACER)	Marion Meiers Juliette Mendelovits
ASER Centre, Pratham India	Rukmini Banerji Shaher Banu Vagh Savitri Bobde
ASER Pakistan	Sehar Saeed
Concern Worldwide	Karyn Beattie Aine Magee Homayoon Shirzad
Concern Worldwide and University College Dublin	Jenny Hobbs
Creative Associates	Joy du Plessis Fathi El-Ashry Karen Tietjen
Durham University	Christine Merrell Peter Tymms
Education Development Center	Nancy Clark-Chiarelli Nathalie Louge
Instituto para el Desarrollo de la Democracia (IPADE)	Vanessa Castro Cardenal
Juarez and Associates, USAID Lifelong Learning Project	Cristina Perdomo Ana Lucía Morales Sierra Leslie Rosales de Véliz Fernando Rubio
Laboratoire de recherche sur les transformations économiques et sociales (LARTES), Jàngandoo	Binta Aw Sall Abdou Aziz Mbodj Diéry Ba Same Bousso Meissa Bèye Diadji Niang

Ministry of Education, Guatemala	María José del Valle Catalán
Ministry of Basic and Secondary Education, The Gambia	Momodou Jeng
RTI International	Keely Alexander Margaret M. Dubeck Amber Gove Emily Kochetkova
Save the Children	Ivelina Borisova Amy Jo Dowd Elliott W. Friedlander Lauren Pisani
Twaweza East Africa	Izel Jepchirchir Kipruto John Kabutha Mugo Mary Goretti Nakabugo Lydia Nakhone Nakhone
UNICEF	Manuel Cardoso
University College London, Institute of Education	Monazza Aslam
University of British Columbia	Linda Siegel
University of Oregon	Sylvia Linan-Thompson
University of Oxford	Pei-tseng Jenny Hsieh
The William and Flora Hewlett Foundation	Patricia Scheid Dana Schmidt
Women Educational Researchers of Kenya	Joyce Kinyanjui

Table of contents

Foreword.....	7
Acknowledgments	8
Abbreviations	9
Introduction	11
Education 2030 and data on learning outcomes	13
Overview of oral reading assessments	15
Recommendations	19
Recommendation 1: Develop an assessment plan for comprehensive reform.....	19
Recommendation 2: Collect additional information to understand the context in which teaching and learning take place.....	27
Recommendation 3: Emphasise the relevant skills—be conscious of differences in culture and orthography of the language	30
Recommendation 4: Properly organize the implementation of activities— logistics and monitoring	33
Recommendation 5: Cater the analysis and communication of results to the target audience ...	35
Recommendation 6: Use the data to raise awareness and design interventions aimed at improving teaching and learning	38
Conclusion	43
References	44



Foreword

With the new Sustainable Development Goal (SDG) for education, governments have pledged to ensure that every child is enrolled in school and learning by 2030. The focus in the past on access to school has given way to a clear commitment to deliver on the transformative power of education with an emphasis on learning. Thus, it is no surprise to find that five of the seven education targets highlight learning skills and outcomes of children and adults.

Reading is considered a gateway skill to all other learning. For this reason, governments are increasingly focused on assessing reading among young children—primarily through oral reading assessments, which are no longer restricted to school settings. A growing number of assessment initiatives led by citizens rather than governments are being conducted in households to help fill the gaps in delivering quality education. While there is strong and systematic support from donors for countries to measure oral reading skills, stronger advocacy and better use of resources are needed to improve learning outcomes. Additionally, further development for the generation and use of assessment data to better inform programmes and policies must be encouraged.

In response, the UNESCO Institute for Statistics (UIS) led a collaborative effort among implementers and practitioners to better understand and communicate what works when implementing oral reading assessments and why, within and across countries. The UIS brought together a diverse community of practitioners (including government officials, donors, non-governmental organizations and university researchers) to identify good practices in the design, implementation and use of oral reading assessments through the production of a series of case studies and articles. This report

highlights the major findings and recommendations drawn from the articles presented in the ebook, *Understanding What Works in Oral Reading Assessments*. The recommendations include a set of concrete guidelines to improve the collection and use of oral assessment data based on experience in more than 60 developing countries.

By presenting a range of experiences from a collaborative but technically rigorous perspective, *Understanding What Works in Oral Reading Assessments* is uniquely designed to encourage different stakeholders to learn from each other in ways that enhance capacity, ownership and cultural sensitivity while fostering innovative forms of international collaboration.

As the SDGs become a reality, governments will need more and better data to inform policies, take corrective action and monitor progress. Early detection of learning gaps will be essential to guiding remedial action and securing the ambition of the new goal to ensure that all children are in school and learning. This publication serves as a unified voice from the community of oral reading assessment practitioners, implementers and donors on the importance of early reading skills to ensure learning for all by 2030.



Silvia Montoya
Director, UNESCO Institute for Statistics

Acknowledgments

The production of the *Understanding What Works in Oral Reading Assessments* ebook and recommendations report would not have been possible without the commitment and efforts of the authors, organizations and national governments that participated in this project. The recommendations presented here draw upon the wealth of experiences of participating authors and organizations in implementing oral reading assessments. Each article in the ebook provides critical information on good practices in the design, implementation and use of data in oral reading assessments.

The UNESCO Institute for Statistics (UIS) would like to thank all research partners for their support throughout this venture, as well as colleagues within the Global Partnership for Education and the William and Flora Hewlett Foundation who provided vital support and encouragement.

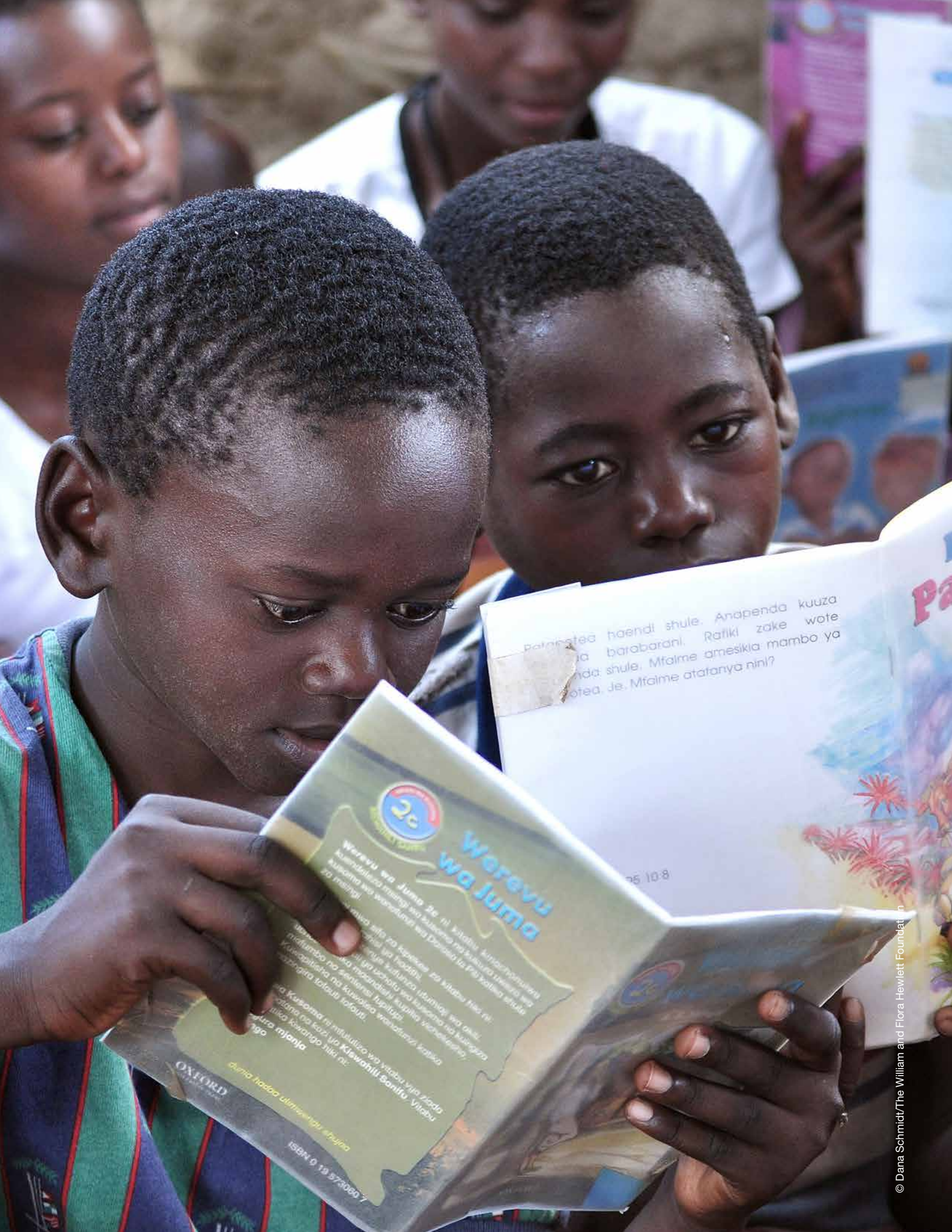
The UIS is grateful to Sylvia Linan-Thompson (University of Oregon) and Linda Siegel (University of British Columbia) for their invaluable input on technical issues.

The UIS thanks all of the authors (see list of contributors) and the peer reviewers for their careful revision: Margaret Dunlop (OISE University of Toronto) and Sheren Hamed (Jordan Education Initiative).

The UIS would also like to thank Maria Elena Brenlla, Nathalie Louge, Sara Ruto, Patricia Scheid and Hannah-May Wilson, who reviewed several articles in the ebook; and Penelope Bender, Luis Crouch and Abbie Raikes for reviewing the recommendations report.

Abbreviations

ASER	Annual Status of Education Report
CETT	Centers for Excellence in Teacher Training
ELGI	<i>Evaluación de Lectura en Grados Iniciales</i> (Reading Assessment for Initial Grades)
GARA	Group Administered Reading Assessment
HLE	Home Literacy Environment
IDELA	International Development and Early Learning Assessment
iPIPS	International Performance Indicators in Primary School
IRT	Item Response Theory
LLECE	<i>Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación</i>
MIA	<i>Medición Independiente de Aprendizajes</i>
MDG	Millennium Development Goals
MICS	Multiple Indicator Cluster Survey
NGO	non-governmental organization
PAL	People's Action for Learning
PASEC	<i>Programme d'Analyse des Systèmes Educatifs de la CONFEMEN</i>
PILNA	Pacific Islands Literacy and Numeracy Assessment
PIPS	Performance Indicators in Primary Schools
PIRLS	Progress in International Reading Literacy Study
RTI	Research Triangle Institute
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SCOPE-Literacy	Standards-based Classroom Observation Protocol for Educators in Literacy
SDG	Sustainable Development Goal
TIMSS	Trends in International Mathematics and Science Study
UIS	UNESCO Institute for Statistics
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WERK	Women Educational Researchers of Kenya
YEGRA	Yemen Early Grade Reading Approach



Palmstee haendi shule. Anapenda kuza
ya barabarani. Rafiki zake wote
nda: shule. Mfalme amesikia mambo ya
wote. Je, Mfalme atatonya nini?

Werevu wa Juma

20
YEARS
ANNIVERSARY

Werevu wa Juma ni kitabu kinafanywa
kuendelea mwingi wa kusoma na kufunza
kusema wa wanaumia wa Darasa la Pili katika shule
za mwingi.

Wanaumia wote wa kwanza za kitabu hii ni:
Kusoma, kufunza, kutambua, kutambua, kutambua
na kutambua. Wanaumia wote wa kitabu hii ni:
Kusoma, kufunza, kutambua, kutambua, kutambua
na kutambua. Wanaumia wote wa kitabu hii ni:
Kusoma, kufunza, kutambua, kutambua, kutambua
na kutambua. Wanaumia wote wa kitabu hii ni:
Kusoma, kufunza, kutambua, kutambua, kutambua
na kutambua.

OXFORD
UNIVERSITY PRESS

Kitabu hiki linawagusa shujaa

ISBN 0 19 577000 7

Introduction

This publication presents a set of recommendations for selecting, conducting and using oral reading assessments. The recommendations highlight basic principles that should be applied in the different stages of executing effective oral reading assessments—from planning and design to implementation and use of the resulting data.

The recommendations are drawn from a series of articles published in the ebook *Understanding What Works in Oral Reading Assessments*. The articles were authored by individuals from organizations that implement assessments, non-governmental organizations (NGOs), academics, ministries of education, donors, international organizations and civil society. In its entirety, the contributions cover experiences from more than 60 developing countries. The ebook featuring this collection of articles can be found on the UIS website at:

 <http://www.uis.unesco.org>

THE SHIFT IN EDUCATIONAL REFORM

Over the last decades, much progress has been made toward ensuring that all children have access to quality education. Despite this progress, considerable challenges remain: 124 million children and youth are out of school (UIS database, 2016) and many more millions of children who are in school are not learning. Research studies and results from learning assessments have exposed the causes of educational failure. These include untrained teachers and absenteeism; mismatches between the language of instruction and children's mother tongue; grade repetition and dropout; children who were never enrolled in school; malnutrition; and more (Sillers, 2015). In many developing countries, a large number of children never start school or drop out, while many of those

who do complete their primary education and graduate do so without acquiring the basic skills required to function in society.

In the last 15 years, the focus of educational reform has been gradually shifting from increasing school attendance to improving the quality of education. The shift in focus to instructional quality has been driven in large part by learning assessment results. Although large-scale international and regional assessments have demonstrated for years that children in developing countries were not learning at the same rate as their counterparts in Western countries, the recent move to assess reading skills in primary school has helped mobilise reform efforts. Since 2009, the number of countries around the world that have collected assessment data to measure early reading skills has increased exponentially through assessments with non-representative sample sizes (studies, impact evaluations, project benchmarks) and those administered at the system-level (examinations, participation in regional or cross-national initiatives and implementing a national learning assessment).

INTRODUCING ORAL ASSESSMENTS

Although there are many types of learning assessments, this report focuses on standardised measures that are designed, administered and scored in a consistent manner and are criterion referenced. In essence, they measure what children are expected to know and be able to do. The assessments are individually administered one child at a time and are direct assessments of foundational skills for learning. We refer to them as oral assessments because children respond orally—usually to written stimuli. Administering an assessment orally is more inclusive as this method



© Dana Schmitt/The William and Flora Hewlett Foundation

allows all children to participate—even those who are not literate. Governments do not necessarily organize the administration of the assessments; generally, there are many partners involved in the different stages of the assessment process. Although the assessments are not explicitly based on the education curriculum in particular countries, they are often compatible with the curriculum as they measure key components of reading and/or numeracy skills acquisition. This report focuses on oral reading assessments.

The use of oral assessments to measure children’s reading development has been instrumental in shifting the focus of educational reform to one that emphasises system accountability, improved instruction and the identification of student learning needs. Unlike international (e.g. PIRLS) and regional assessments (LLECE, PASEC, PILNA, SACMEQ), oral assessments can be—relative to policy impact—smaller, quicker and cheaper (Wagner, 2011) to design and administer in local languages. These are critical features in settings where children enter school speaking a number of different

languages and funds for conducting assessments may be limited. Further, results are actionable, targeted to early reading and are usually available for dissemination in a shorter timeframe compared to regional or international assessments. It is these last three characteristics that have contributed to the impetus needed to change the focus of educational reform from access to education to quality of instruction and student learning outcomes.

It is important, however, to recognise the limitations of oral reading assessments. First, they are resource intensive in terms of staff required to complete the process. Second, they are time consuming as they involve training several groups of individuals to perform the various tasks required. Third, the reading comprehension measures are limited and may not discriminate among students for several reasons: there are few items; the test generally allows lookbacks; and the questions included are typically explicit and inferential so do not involve interpreting, integrating ideas and information, or evaluating and critiquing content.

Education 2030 and data on learning outcomes

With the adoption of the Sustainable Development Goals (SDGs), governments have pledged to ensure that every child is enrolled in school and learning by 2030. The focus on learning outcomes is a shift from the Millennium Development Goals (MDGs), which focused on ensuring access to, participation in and completion of formal primary education (UNESCO-TAG, 2015).

Policymakers at the global and national levels clearly recognise the importance of determining whether the quality of education is improving and the role that the monitoring of learning outcomes plays in achieving this end. It is not enough to know how many children are enrolled in school or how many teachers are hired to reach the SDGs. They need to know whether children possess the basic reading and mathematics skills essential to future learning. They need to know what children can and cannot do early on to ensure that there are policies and practices in place to support early intervention and remediation. Waiting until the end of primary education to ascertain learning levels will be too late for many children.

To help transform this promise into action, governments will need more and better data to identify areas of improvement, install change and monitor progress. The good news is that through household surveys, learning assessments and research studies, educators, administrators and other stakeholders have been engaged in answering questions, such as: What are children learning? Where are they learning? And who is being left behind?

The ability to read is essential for progress in the education system. Having relevant, high-quality early grade literacy data is a crucial step in attaining

this goal. Although assessment is vital to guiding government policy and changes to instruction, it alone is not enough. Data should be analysed and governments should continuously evaluate their policy agendas, school-level implementation and progress through the use of assessments and their results to ensure that all children are learning.

A FOCUS ON READING

The SDG for education calls for monitoring learning outcomes, and several indicators in the Education 2030 Framework for Action specifically refer to reading. Reading is considered a gateway skill to all other learning. Children who fail to develop appropriate reading skills in the first few years of schooling are likely to continue to lag behind their peers (Juel, 1988). In low income countries, these children often drop out of school before completing primary education. Thus, ensuring that all children learn to read has served as the impetus for assessing reading in the early years of schooling—primarily, through oral reading assessments. Although there is consensus that reading is an important skill, there is, however, less agreement on what skills should be assessed and how they should be assessed.

SHARING EXPERIENCES TO UNDERSTAND WHAT WORKS IN ORAL READING ASSESSMENTS

Given the focus on reading and on trying to guarantee early success as a contribution to primary school completion, many organizations have started using one-on-one oral assessments that involve printed stimuli. The rationale for using oral assessments as opposed to written assessments will be described throughout this report. Of these, a few warrant mentioning at the outset.

First, participation in most pencil and paper assessments requires some word reading ability so if many children are not able to respond, there will be very low discriminant capacity at the lower end of the scale. Also, given the relative informality of many school settings, it is possible that in group assessments, especially if teachers are present or the assessment content is leaked, children may be coached or even helped during a group-administered, pencil-and-paper test. Assessments that are orally administered, one-on-one, by individuals who are from outside the school, help circumvent some of these problems. In addition, oral assessments can assess very basic oral skills such as phonological awareness and basic literacy skills, such as letter knowledge.

For these reasons, the use of oral assessments has become relatively widespread. Despite some commonalities among the instruments used, there are also differences in the purpose, design and administration of these assessments. Given the wide array of assessments available to practitioners, the UNESCO Institute for Statistics (UIS) led a collaborative effort with organizations that have been actively financing, designing and implementing oral assessments (see **Box 1**). Representatives from these organizations were asked to submit case studies and position papers that exemplify good practices. The information from these papers was then synthesised and used to derive the resulting recommendations.

It is hoped that these recommendations will provide the field with a set of concrete guidelines to improve data collection and their use.

Box 1. Collaborators of *Understanding What Works in Oral Reading Assessments*

- 30 organizations
- 50 contributors
- Combined experiences from more than 60 developing countries

The methodology of this collaborative exercise drew on the following principles:

1. *Moving towards consensus*. Being a consensus-building exercise, the organizations' own know-how served as the starting point. Experiences were shared and different perspectives were compared.
2. *Focus on identifying balance between cultural specificity and global applicability*. Maintaining equilibrium between these two principles and addressing the challenge of identifying the culturally specific lessons that apply only to certain regional, linguistic or cultural contexts was deemed important. Equally important is the goal to identify overall principles that may apply to a wide variety of developing contexts.
3. *Parsimony*. It was key to emphasise the importance of streamlining and simplifying assessment instruments and methodologies without incurring a loss of precision and explanatory power as these are relevant to policymaking.

The 20-month process that culminated in the development of these recommendations can be summarised in **Figure 1**.

Figure 1. Development phases of the oral reading assessments recommendations



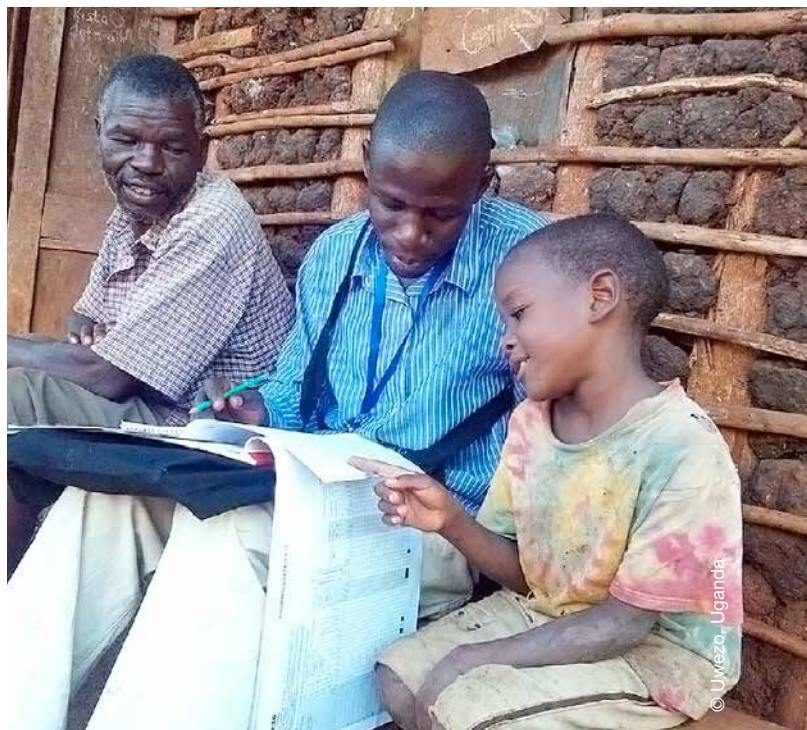
Overview of oral reading assessments

Assessment, in educational contexts, refers to a variety of methods and tools that can be used to evaluate, measure and document learning progress and skill acquisition (see **Box 2**). In addition to providing information on current student achievement, empirical data can be used to determine teaching instruction quality, identify students' learning needs or evaluate language ability. The most common use of oral reading assessments is to determine students' current level of performance. These data often serve as a baseline for specific interventions or generalised reform efforts.

Box 2. Commonalities among oral reading assessments

Although oral reading assessments are designed for different purposes, they share some characteristics. Any given assessment is typically a standardised measure that is designed, administered and scored in a consistent manner and is criterion referenced. The assessments measure what children are expected to know and be able to do. They are individually administered, direct assessments of key components of reading skills acquisition. Most often, these are assessments of learning (i.e. they are designed to inform stakeholders and not teachers).

Once a need for reform has been established and an intervention is implemented, oral assessments can serve as an outcome measure to determine the effect of the intervention. When assessments are used to determine the effect of an intervention, it serves as an evaluation tool. According to Fenton (1996), 'evaluation is the application of a standard and a decision-making system to assessment data to produce judgments about the amount and adequacy of the learning that has taken place'. Essential to this process is the availability of



standard or normative scores that provide parents, educators, administrators and donors with an index by which to judge whether learning progress is meaningful. This section will provide an overview of the different types of oral assessments.

ACCOUNTABILITY ASSESSMENTS

Accountability assessments are used to report to the public and other stakeholders on educational trends and to demonstrate the effectiveness of the education system in serving children and in meeting the needs of the community and state.

Citizen-led assessments

Citizen-led assessments are generally those that are led by citizens or civil society organizations

TABLE 1

Citizen-led assessments

Citizen-led assessment	Country	Target population (assessed children)	Year initiative was launched
ASER	India	5–16 years old	2005
ASER	Pakistan	5–16 years old	2008
Beekunko	Mali	6–14 years old	2011
Jàngandoo	Senegal	6–14 years old	2012
Uwezo	Kenya	6–16 years old	2009
Uwezo	Uganda	6–16 years old	2009
Uwezo	United Republic of Tanzania	7–16 years old	2009

Note: Other citizen-led assessments include Medición Independiente de Aprendizajes (MIA) launched in Mexico in 2014 and LEARNigeria in Nigeria in 2015; the assessments target children aged 5–16 years and 5–15 years, respectively. LEARNigeria, similarly to ASER-India, also surveys all children aged 3–15 years yet only those aged 5 or older are assessed. Both MIA and LEARNigeria are not yet administered to a nationally-representative sample of children.

Source: adapted from (Aslam et al., 2016) and the UIS Catalogue of Learning Assessments, 2016

rather than by governments (see **Table 1**). They are conducted in households rather than in schools and measure basic reading and numeracy skills. Citizen-led assessments can provide recurrent estimates of children’s basic learning levels and (so far) tend to be similar in design and administration. Citizen-led assessments are a different model of assessment. Rather than being in the hands of a limited number of professionals, the community has a stake in administering and interpreting these assessments.

Volunteers administer the measurement tools that assess children’s reading skills in homes or communities. Children’s reading levels are typically characterised as being either at the letter, word or passage level (often two passages with varying levels of difficulty are included in the assessment). This approach allows stakeholders to track changes in the number of students at each level over time.

The results from citizen-led assessments are used for accountability and advocacy by (Aslam et al., 2016):

- generating nationally representative and locally owned data on acquisition of foundational skills that are not dependent on school access;
- helping re-orient the debate from school access to improved learning for all;
- creating new opportunities for citizens to better understand the status of their children’s learning

so that they can decide for themselves whether governments are delivering on promises related to equity and quality in education;

- promoting new mechanisms for evidence-based policy, proven programme interventions and actions to improve learning.
- creating a sense of community and shared purpose.

There are two further points worth noting about citizen-led assessments. First, while citizen-led assessments have mostly been used for generating accountability pressure, it typically has not been a high-stakes accountability pressure tied to particular teachers or schools. Rather, their main purpose has usually focused on education-system accountability or overall community-based accountability. In addition, they have also been used in the classroom to group children by skill and to place them at the right level, rather than based on grade or age or curricular expectations. The approach of teaching at the right level is currently gaining some traction among educators in developing countries.

School-based assessments

A second type of accountability assessment is the school-based oral assessment. The most commonly used is the Early Grade Reading Assessment, which has also been used in settings other than in schools. Other widely used school administered

assessments include the Reading Assessment for Initial Grades (ELGI in Spanish) and Literacy Boost. These assessments are administered in schools and results are often used to advocate for educational reform. In the reform process, stakeholders use data from these assessments to make decisions on the use and effectiveness of resources, personnel and institutions. Reform efforts initiated after data have been collected on a national sample often include changes in instructional approaches and curriculum, textbook development and resource allocation. Although one could classify these assessments as driving accountability, it is important to note that the accountability sought here is at the level of the teacher support system, the system that provides learning materials to learners and the overall policy. Few, if any, of these assessments are used to assess individual teachers and as a matter of fact, they are designed to be sample-based assessments that do not identify individual teachers or learners.

Since literacy begins before formal schooling, assessments, such as the International Development and Early Learning Assessment (IDELA) (see article by Dowd et al.) and the Performance Indicators in Primary School (PIPS) (see article by Merrel and Tymms), seek to identify which skills children possess prior to beginning formal primary education. Like the measures used with school-aged children,

results from these assessments provide data on children's level of skill acquisition and can be used to improve early childhood programmes. Measures designed to assess children's knowledge and skills at school entry can also provide Grade 1 teachers with information on children's relative learning performance that can be used to plan instruction to support all learners.

School-based oral reading assessments have also been used as outcome measures in the evaluation of intervention projects in a variety of contexts. The data are collected at two or three points during the span of a project. When used as a formative measure, students are assessed while the intervention is being implemented and results are used to make programmatic changes. The use of data to make decisions is critical when implementing a new instructional approach. However, at the end of the project, results of the summative assessment are used to determine the effect of the intervention or reform effort. Literacy Boost, for instance, a well-known reading intervention, has been implemented in a number of countries and in 35 languages. Results from the Literacy Boost assessments are used to shape and evaluate the implementation of Literacy Boost programmes.

où est mon chaton?
où est mon chaton?
Dit maman chat.
Je le cherche partout.
Pas chez moi, dit le chien
pas chez moi, dit la chèvre
Pas chez moi, dit la vache
Ah! le voilà!
chaton, bien o ud
Se cache dans les choux

Mbare

SIDA
Protégez-vous!

Recommendations

The following section presents the recommendations for selecting, planning, implementing and using oral reading assessments. The recommendations highlight basic principles that should be applied in the different stages of such assessments and are based on the articles presented in the ebook *Understanding What Works in Oral Reading Assessments*. Although the articles explore a wide range of oral reading assessments conducted on different scales, the following recommendations pertain to practices that can be scaled-up to the system-level.

RECOMMENDATION 1:

Develop an assessment plan for comprehensive reform

- Ministers of education or educators must make these three decisions when developing an assessment plan:
 - determine the level of assessment or who will be assessed
 - its purpose or why it will be administered
 - the object of assessment or what knowledge, skills, language level, perceptions or attitudes will be assessed.
- When developing an assessment plan, assembling a solid team of partners, ensuring data quality, constructing a vision for what will be done with the results and creating an itemised budget are of critical importance.

Assessment has taken center stage in education reform. Currently, data from oral assessments are used to make system-level programmatic decisions to inform reform efforts or individual projects. If the sustained use of assessments for instructional decision making by ministries of education is the goal, then the current use of early grade assessment measures needs to be expanded and at the same time aligned with the rest of the assessment systems and frameworks within countries.

At the national level, an assessment plan is needed for comprehensive reform. An assessment plan outlines: what data will be collected, by whom and for what purpose; the process for reviewing data, policies and procedures to guide feedback results; and the process for modifying the programme

or curriculum. Summative assessments evaluate student learning at the end of a specific instructional period. Interim assessments evaluate where students are in their learning progress and determine whether they are on track.

To improve instruction, ministers of education and educators must make decisions on which assessments will help them develop an assessment plan that will provide data that are not only useful but that can be collected with fidelity. They will have to determine the level of assessment, its purpose and the object of assessment.

LEVEL OF ASSESSMENT

Current oral assessments are individually administered. Individual assessments are preferable



© María José del Valle Catalán, Guatemala

when assessing young children, particularly when assessing reading, since children may not be able to read yet and critical skills in assessing reading, such as phonological awareness, are best assessed individually. Alignment across the various assessments is advised to ensure that data provide information on students' progress across time and across measures. Save the Children's IDELA and Literacy Boost assessments provide an example of how measures can be aligned. These assessments measure children's early learning and developing reading skills from age 3 to Grade 3, presenting a continuous assessment framework between pre-primary and primary education. In addition, the range of skills assessed when measures are aligned helps to avoid floor effects. That is, these assessments pay more attention to what is happening at the lower end of the skill distribution by testing a more basic version of the same skills. This makes them particularly well suited to marginalised populations. Further, they promote the inclusion of a range of continuous

indicators, spanning from foundational to higher order skills, and hold the goal of learning to read as the ultimate non-negotiable outcome to measuring comprehension (Dowd et al., 2016).

It should be noted that although individually-administered oral assessments have been the de facto norm, group-administered oral assessments have been used in developing countries. The implementers of the Centers for Excellence in Teacher Training (CETT) (Chesterfield and Abreu-Combs, 2011) developed group-administered oral assessments of early reading skills to measure the impact of the project. More recently, researchers at the Research Triangle Institute (RTI) International have performed psychometric research on the suitability of group-administered oral assessments in developing countries. With the Group Administered Reading Assessment (GARA), the assessor orally administers a reading assessment to a group of students and their responses are collected using paper-and-pencil student response sheets. Since

the assessment tool is in the form of multiple-choice questions, children's writing skills are not being tested along with reading—with the exception of the writing dictation subtask. Like the measures developed by CETT, GARA differs from other group-administered reading assessments (e.g. LLECE, PIRLS, PASEC, SACMEQ) in that the test is not reliant on passage reading ability. Since it begins with skills as simple as letter names/sounds, group-administered assessments cover the same range of skills as measured by the EGRA. The goal of group-administered reading assessments is mostly to lower the cost of training as it is time-consuming to train assessment administrators to conduct oral assessments. The GARA was conducted in Egypt (RTI International, 2014). It is still evolving and being piloted in different countries and contexts.

PURPOSE OF ASSESSMENT

Determining the purpose of the assessment is important due to various implications, such as who will be assessed, how often and where. The plan must also stipulate who will conduct the assessment. Some critical points to consider:

1. What is the purpose of assessment?

This question will drive all other decisions. It is also important to note that the purpose of assessment may change over time. An initial assessment plan may propose including children in the early primary grades to obtain a baseline to determine whether reform is needed nationally or in a specific region. When a ministry of education is implementing a new instructional approach, it may want to assess children more frequently using both formative and summative assessments but only target the children that are part of the intervention. However, if the purpose is accountability and an early grade reading assessment is being integrated into a system-wide assessment plan, then decisions will be driven by curricular factors—for example, the assessment can be administered at the end of the first year of instruction and then again at the end of the first cycle. Another purpose may be to determine the literacy levels of children outside the education system, including

children in informal programmes or those who never attended school or who have dropped out.

2. Who will be assessed?

In the case of school-based assessments, this refers to both the grade levels to be assessed and the number of students who will be assessed. Identifying a representative sample of students is key. When assessing students during an intervention, both intervention and control students should be assessed. Whether assessment is part of a national plan or an evaluation of an intervention, it is important to first identify all groups of children that need to be included (e.g. from different language groups, genders, socio-economic statuses, geographic locations, etc.).

Next, it is key to determine the appropriate proportions for each group. Appropriate representation will ensure that results can be generalised to all students in the education system. These principles also apply to household assessments. In the case of household assessments, the purpose of the assessment may be to determine the literacy levels of preschool-aged children or out-of-school youths. It is also important to find representative samples of children. Although it may be more difficult to identify the eligible population when conducting household assessments, there are formal and informal avenues to collect census information and create sampling frameworks. Official census as well as village, tribal or church registrars can be used to identify participants and to ensure that the sample is representative.

3. How often will children be assessed?

The timing of the assessment is also based on the purpose. For example, when implementing a new reform effort, more frequent assessment may be necessary to monitor implementation. Once the reform is well under way, less frequent assessments may be required or only assessments of students in particular grade levels. Another important aspect of measurement is assessing change and growth in



literacy skills. A measurement of change requires at least two points in time. This repeated assessment may not be feasible due to logistics, cost and/or other factors but it is desirable when possible.

4. Where will children be assessed?

If there is an interest in determining what all school-aged children in a particular context know, then a household-based assessment may be more useful, particularly in a context where not all children attend school. However, if the purpose is to determine how well children in school are learning and/or if all children are in school, then a school-based assessment presents a better option. This assessment could be combined with a targeted assessment of children not in school with a specialised sampling approach to capture data on them. An education system may require information from both school-based and household-based assessments to determine the literacy landscape in the country. However, assessing at both levels may not be possible or sustainable.

5. Who will collect the data and how?

Assessment data can be collected by trained community volunteers, paid trained assessors or

ministry of education staff, including teachers. Determining who will assess students will depend on resource availability and capacity. To ensure that the data are collected reliably, the assessors should be fluent in the language of the assessment and should receive adequate training, including (when possible) training on site at the local schools. Additionally, assessors should be assessed before the data collection to ensure that they are well prepared. Measuring inter-rater reliability is highly encouraged to determine assessor readiness.

Oral assessment data can be collected using paper or digital means. Determining which data collection method to use will depend on resources. Capturing the data using digital means will depend on the daily access to electricity to recharge tablets or phones and online connectivity to upload data. Some of the general advantages to collecting data electronically include: rapid availability of assessment data; improved data accuracy and fewer measurement errors due to missing fields, data transcription errors, invalid data types or formats and illegible or incomprehensible data; a reduced amount of paper and supplies required; as well as more simplified logistics to prepare and manage the data collected compared to paper assessments (i.e. no photocopying, sorting, stapling, packaging, etc.). Rapid availability of the data also makes supervision easier and can result in immediate correction of field problems (e.g. quicker online support to an assessor who encounters problems).

OBJECT OF ASSESSMENT

The object of assessment refers to what is assessed. It is important to identify what domains and constructs will be assessed to determine children's knowledge and skills. The focus of oral assessments has been on early reading skills. There appears to be consensus on what to measure in both foundational and higher order skills based on the substantial literature in the field of reading development.

Box 3. Adapting an existing instrument or designing a new assessment tool

Valid and reliable instruments and tools that have been developed for various purposes can be adapted to different contexts. For example, the Women Educational Researchers of Kenya (WERK) has developed an oral reading assessment for the Maa language based on the EGRA and Uwezo Kiswahili assessment tools (Kinyanjui, 2016). Adapting instruments to each new context requires knowledge of the linguistic structure of the language among the students to be assessed, the context and often some sense of curricular expectations as well as the availability of reading or language textbooks. Only when adapted correctly and applied using proper assessment techniques will the results yield a reliable and valid depiction of skills (Dubeck et al., 2016). Note that when adapting existing tools, piloting is critical. The *EGRA Toolkit* provides detailed guidance on how to develop and adapt an EGRA.

If designing a new oral reading assessment is the goal, then assessment developers must address the following factors:

- Testing economy: how much information do you get from the battery of tests? How many reading constructs will you assess?
- Efficiency and predictive validity: how much information do you get for the effort? Limit the measure to those that are most predictive. Assessments should take no longer than 15-30 minutes.
- Task difficulty: which skills and knowledge will you measure? Are they appropriate for the language and the reading level of the students?
- Developmental validity: how well will the items hold up over time? Have you avoided floor and ceiling effects?

To ensure assessments are reliable, they must be developed through a rigorous process. Overall the tasks to be included in the assessment should be:

- Research-based and capable of assessing critical aspects of literacy;
- Built around contexts likely to be familiar to students in the early years of school;
- Able to be administered by the student's own teacher (when being used by teachers and schools for formative purposes as opposed to outside assessors). In this case, the tasks should be easy for teachers to administer and should be supported with clear and explicit marking and recording guides (Meiers and Mendelovits, 2016).

Although oral reading assessments do not vary much in length of administration, they vary in the type of data they provide. For example, while all assessments provide data on students' alphabet knowledge, word-level reading and text reading, the type of information varies. Some instruments assess students on all letters while others assess students on the easiest and/or most difficult. Some assessments are timed while others are not. As a result, these assessments provide different levels or types of data with regard to children's reading ability that range from basic categories (i.e. letter, word or text) to fluency rates on a number of skills.

There are also differences in the number of constructs that are assessed. If the purpose of the

assessment is to determine what reading skills and knowledge children possess and at what point they developed these, it is helpful to assess foundational skills that are predictive of later reading abilities and higher order skills when planning an intervention. Although not all reading skills are amenable to quick assessments, the data from even brief assessments can provide an index of what students know and can inform reform efforts.

Figure 2 illustrates the steps in planning the EGRA assessment. The timeline is to be used for planning purposes.

Figure 2. The Early Grade Reading Assessment Timeline



Source: Kochetkova and Dubeck, 2016

IMPORTANT PRACTICAL ELEMENTS TO CONSIDER IN THE ASSESSMENT PLAN

As noted above, an assessment plan outlines what data will be collected, by whom and for what purpose; the process for reviewing data, policies and procedures to guide feedback results; and the process for modifying the programme or curriculum. Aside from the key structural components for developing an assessment plan, there are also two important practical elements to consider: building a good team of collaborators and budget planning.¹

1. A solid team of partners

Strategic partnerships are critical for sharing knowledge and increasing ownership of the assessment results. Partners include donors, ministry staff and technical collaborators. The assessment plan must be shared with the various partners and should be aligned with the country's priorities. It is strongly advised that ministry staff be engaged at all levels of the assessment. Experience shows that those officials who participate in the development and implementation of the assessment will understand its applicability to their national context and are able to advocate for its use when required. Involving local actors also ensures that their skills are built to carry out other assessment activities in the future. Sometimes involving high-level officials in the field work—even for a day—can also prove useful as they can develop an immediate and practical sense of how the assessment works and gain firsthand a sense of children's reading levels.

2. Budget planning

Budget planning depends on how much work has already been done; for example, there are differences in the initial cost of an assessment and a reapplication. In general, costs of oral reading

¹ The recommendations in this subsection have been compiled by the UIS based on a questionnaire circulated among a select number of authors. The questionnaire focused on experiences and lessons learned related specifically to budgeting and assessment planning.



assessments will vary by country and are dependent on sample size, level and number of disaggregation desired, local inputs of labor and transportation and the use of technology for data collection. Of the assessment experiences included in the ebook, the cost per learner ranged from less than a dollar (USD) to a few hundred dollars. When contributors to the ebook were asked to categorise the proportion of funds allocated to each of the assessment stages, almost all allocated the largest proportion to test application (implementation) and the lowest to dissemination activities. The assessment cost categories were based on a breakdown for budgeting proposed by Wagner et al. (2011) which included test preparation, test application (implementation), processing and analysis, dissemination, and institutional costs.

When planning the initial assessment, donors and non-governmental organizations (NGOs) are encouraged to involve in-country teams who can provide more accurate local costs. It is also prudent to leave a margin in the budget for unexpected consequences/missed expenses. This will ensure that there are sufficient funds for each stage of the process and avoid delays or shortcuts due to insufficient funds.

Further, good planning for the first assessment can result in reduced costs for future implementations. Taking the time to plan carefully and develop a valid and reliable assessment will help avoid incurring unexpected costs later on. It will ensure that results are valid—the assessment will stand the test of time, reducing the likelihood that it will have to be modified later which will result in additional costs and loss of comparability over time. If there is a desire to measure change over time, then multiple equated assessment forms must be designed from the outset (more information is available in the *EGRA Toolkit*). Careful planning will also

reduce the likelihood that training materials will have to be modified down the road. While there are ways to reduce costs, there are areas that are non-negotiable, such as ensuring the assessment is properly developed and piloted during test preparation. Other areas, such as who assesses are negotiable—one way to reduce assessor costs is to train ministry staff or salaried officials as assessors or to use volunteers. Another way to reduce costs is to take advantage of pre-planned activities to disseminate the results of the assessment.

It is advisable to construct a strategy to enhance national capacity in implementing assessments. It should encompass a plan to support ministries in providing oversight and management of the assessment. If strengthening national capacities in analyzing and interpreting data is desired, then planning for activities such as theoretical training sessions on measurement method models, use of special statistical software packages and data analysis techniques should be considered.



RECOMMENDATION 2:

Collect additional information to understand the context in which teaching and learning take place

- Data on reading achievement alone are not sufficient to design sound interventions. Additional information is necessary to ensure understanding of the assessment results.
- Choosing what additional information to collect is dependent on the purpose of the assessment and the context. The additional variables can be related but are not limited to: language(s), home literacy practices and quality of instruction.

Understanding the social, economic and political settings of schooling as well as the educational environment in the country in question contributes to appreciating how those contexts impact learning and is critical when interpreting assessment data or designing interventions. Although oral assessments of early reading provide useful data on children's current level of performance, learning occurs in a variety of contexts. Therefore, in addition to decisions related to the type of assessment to be implemented (benchmark, evaluation, national diagnostic, etc.), countries should also collect relevant contextual information to ensure understanding of the assessment results—a necessary step to informing and formulating targeted interventions/policy. The decision on which additional variables to collect is tied to the purpose of the assessment and the context in which teaching and learning take place.

The measurement of the context of early reading skills is complex and there are many variables to consider. These variables include but are not limited to: national language(s); local language(s); language(s) of instruction; instructional practices; organization of the school system and access to formal schooling; teacher training; curriculum; exposure to oral and written languages; funding available for education; the availability of teaching materials and books; gender issues in terms of access to education; home and environmental practices related to literacy (e.g. presence of reading materials in the home, print-richness of the environment); and socio-economic and cultural conditions. It is important to be aware of these variables when assessing the reading

skills of an area or country. In some countries or certain communities in rural areas, access to water, sanitation, healthcare and education is quite limited or non-existent. Such variables should be systematically recorded in order to provide a clear picture of the context.

LINGUISTIC CONTEXT

In most countries, children are growing and learning in bilingual or multilingual environments. Children's proficiency in the language of instruction can have an impact on assessment and the effect of interventions. Recently, the use of oral assessments has expanded from assessing reading and mathematics to the assessment of children's oral language proficiency. Although there has been a concerted effort to educate children in their first language or mother tongue, the reality is that many children in low-income countries are learning to read in a second or third language—even if this runs counter to a national policy that the language of instruction in the first few grades should be the children's first or home language. When children are receiving instruction in their first language, they often come to school with very limited oral language skills (Hart and Risely, 2003).

In these contexts, information on children's language proficiency in their first language and the language of instruction is essential for programme planning and to ensure that children are provided instruction that promotes learning success by supporting language acquisition in the language of instruction. The development of measures that provide educators with information on children's

language profile is just beginning. Processes for developing language proficiency measures have been developed in Latin America in Spanish and in indigenous languages in Guatemala (Rosales de Véliz et al., 2016). However, replication in other languages and contexts is needed. If possible, it would be desirable to assess children in all the languages that they speak. This type of multilingual assessment (if relevant) will provide a clearer picture of children's language skills.

QUALITY OF INSTRUCTION

Another area to consider when developing an assessment plan is the quality of instruction. Although collecting data on the fidelity of implementation is essential when introducing an intervention or a reform effort, more general descriptions of instruction are useful when interpreting assessment results.

There is a significant link between teaching quality and student achievement. If possible and when appropriate, teaching quality should be assessed. There are a number of dimensions on which to assess teacher quality. Some of these include the teacher's use of instructional time, instructional strategies, the materials used and the student grouping strategies (individual, small group, entire class). The Standards-based Classroom Observation Protocol for Educators in Literacy (SCOPE-Literacy) was developed to provide information on the quality of instruction. The measure has two dimensions: classroom structure as well as language and literacy instruction. Both areas contribute to quality instruction (Yoon et al, 2007). The results from the SCOPE-Literacy can be used to target professional development and improve instruction (Clark-Chiarelli and Louge, 2016).

CHILDREN'S HOME ENVIRONMENT

In developing contexts, understanding the literacy environment in the home can help explain reading achievement. Measuring the home literacy environment requires collecting and analysing data on: the value placed on reading and the drive to achieve; the availability of reading materials; the frequency of reading to and by children; and opportunities for verbal interaction (Hess and Holloway, 1984). Save the Children further claims that children's motivation and opportunities to read inside and outside both the home and the school should also be considered (Dowd and Friedlander, 2016). Including these elements will enable a better understanding and a broader evidence base that more appropriately represents the rich variety of learning environments in different languages, cultures, physical environments and overall living situations that exist throughout the world (see **Box 4**). Measuring the home literacy environment is done through a survey of family members and activities as well as follow-up questions on the motivation for reading and literacy use outside the home.

Save the Children's Home Literacy Environment (HLE) survey complements the data collected in schools by providing data on variables associated with academic success, such as opportunities for verbal interactions or the availability of reading materials at home (see **Figure 3**). Data can be used to explain the differential effects of interventions which enable implementers (NGOs or ministries of education) to make decisions on how to improve or adapt their programmes for different groups of children. For example, if intervention results show that children who have fewer opportunities to read outside of school have lower reading scores, educators can consider ways to collaborate with other stakeholders to increase those children's opportunities to read both in and out of school.

Box 4. Complex environments

For many years, the international community has supported education in complex contexts albeit mainly through building and rehabilitating infrastructure and providing school supplies and teachers—interventions that improve access to education. However, surmounting the educational barriers found in politically complex environments requires significantly greater initiatives than simply improving access. In reality, the challenges in these environments include overlapping barriers to learning, such as poverty, conflict, gender inequality, low exposure to print, illiteracy of parents and food insecurity. Thus, new approaches to improve learning outcomes are much needed and learning assessments can provide a clearer picture to better inform reform efforts.

Experiences from Somalia and Afghanistan show that there is value in implementing oral reading assessments as a first (and difficult) step in a long process to improve learning. While the administration of the assessment may have been fraught with difficulties due to transport, weather and security problems, the countries have found the efforts to be worthwhile.

In complex situations, there will inevitably be some compromises. However, until assessments are being administered in these contexts, some of the poorest and most vulnerable children will inevitably be left behind. Thus, even small data sets may be useful. Compromises on sample sizes or on supervision of the test administration is acceptable in these cases—although the compromises should be ‘principled’ (i.e. the limitations they impose should be acknowledged and should be accounted for when the data are reported). When compromises are ‘principled,’ the claims that are made based on the data have to be explicitly cautious.

Source: adapted from (Shizad and Magee, 2016) and (Beattie and Hobbs, 2016)

Figure 3. HLE survey matrix

Name/initials	Relationship 1=Mom, 2=Dad, 3=Sister, 4=Brother, 5=Grandma, 6=Grandpa, 7=Other Female, 8=Other Male	Seen reading 1=YES, 0=NO	Told/helped you to study 1=YES, 0=NO	Read to you 1=YES, 0=NO	Told you a story 1=YES, 0=NO
Other than at school, did anyone outside your home read to you last week?				__No (0)	__Yes (1)
Other than school, did you read to anyone outside your home last week?				__No (0)	__Yes (1)
Other than at school, did you read alone last week?				__No (0)	__Yes (1)
In the last week, did you use your reading skills outside of school?				__No (0)	__Yes (1)
If yes, where? _____					__Yes (1)
In the last week, have you helped anyone using your reading skills?				__No (0)	__Yes (1)

Source: Dowd and Friedlander, 2016

RECOMMENDATION 3:

Emphasise the relevant skills—be conscious of differences in culture and orthography of the language

- All children should know the names of the letters (in alphabetic languages), be able to read words and pseudo words.
- Across languages, fluent reading contributes to reading comprehension. Yet, when assessing children, it is important to remember that the relative importance of speed and accuracy is dependent on the orthography and culture. Speed should not be pursued for its own sake.
- Comprehension is the ultimate goal of reading and it must be measured, even if done using a limited number of tasks.
- Most oral reading assessments are not designed to be comparable across countries or cultures. However, research shows that there are some skills and areas of development that can be compared.

Reading develops in similar ways across languages and settings. ‘Brains are basically the same across languages. Orthographies are not’ (Goswami, 2006). Although children require similar skills to become proficient readers, the skills that must be emphasised will vary according to the orthography of the language. The decision to measure specific skills depends on language, script, orthography and instructional methodologies.

ASSESSMENT COMPONENTS

Oral reading assessments generally include basic constructs, such as letter recognition, phonological awareness, word reading and pseudo-word reading as these are the foundations for pre-reading skills and higher order skills (i.e. vocabulary, oral reading fluency, comprehension, etc.). Oral reading assessments that are appropriate for early grades and that vary in the constructs they assess have been developed and implemented in more than 100 languages.

1. Phonological awareness

When determining which constructs to include, consider the role they play in the target language. Certain skills are important precursors to the development of reading. These skills are alphabet knowledge and phonological awareness. Phonological awareness is the ability to hear the sounds within words and to manipulate these

sounds. An example of a phonological awareness task is asking a child to say the word ‘book’ without the /b/ sound and then asking the child to say the sound of the missing letter rather than the name of the letter. Phonological awareness skills will continue to develop as children learn to read.

2. Word reading

Another important precursor to the development of reading skills is the knowledge of letters and their sounds. Across languages, children have to recognise the symbols or graphemes used in the language and to link sounds to the graphemes. This is the basis for decoding words. Although children can learn to decode once they know a few graphemes and their sounds, the total number of symbols that children have to learn will affect how long it takes them to become proficient readers. When developing assessments, select the graphemes that are grade appropriate.

One of the most fundamental skills that should be measured is reading single isolated words when there are no text clues to the meaning of the word. Children need repeated exposure to words and text to develop automaticity in reading. The goal is for all word reading to eventually become automatic. As children get older, and the words they are expected to read are also longer or more morphologically complex, grouping larger units together is a more efficient way to read. Therefore,

the ages of the children who are being assessed informs which types of words should be included in an assessment. Automatic word recognition is necessary for fluent passage reading. Oral assessments of reading should be based, at least in part, on phonics (i.e. knowing the sounds of the letters). Skills in phonics can be assessed by the reading of non-words (technically called pseudo words) that are pronounceable combinations of letters or characters. This task tests the ability to decode print into the sounds of the language. This phonological processing ability is important for decoding new words and names that have not been previously encountered.

3. Fluency

Fluency refers to reading accurately with adequate speed and prosody. Across languages, fluent reading contributes to reading comprehension. Yet, when assessing students, it is important to remember that the relative importance of speed and accuracy is dependent on the orthography of the language. In transparent orthographies, speed is a more important indicator of reading skill but in opaque orthographies, accuracy is a better indicator. This is because in opaque orthographies, the sound/grapheme relationships are less predictable and reading words incorrectly can have an impact on comprehension. One challenging issue is whether or not to measure the speed of reading. Although the ability to read quickly enough to process and store information is important, reading speed can be difficult to measure with any precision in the field. Oral reading fluency is important and often assessed but finding the proper metric that would be comparable across languages is difficult.

4. Comprehension

Comprehension is the most complex of the reading skills and represents the ultimate goal in reading. To comprehend what is read requires word-level skills; vocabulary knowledge; oral language skills; reading with a modicum of fluency; broad conceptual knowledge, thinking and reasoning skills; and specific reading comprehension strategies. There

are several reasons why comprehension is difficult to assess well when using a brief measure. In some cases, children may not have the basic reading skills needed to make measuring comprehension feasible. In other instances, there may not be enough items to sufficiently assess comprehension. Although oral reading fluency is correlated to reading comprehension, the correlation can be influenced by factors, such as students reading in a second language in which they are not yet proficient.

Assessments that include a variety of tasks may provide better—but not perfect—clues as to why children have trouble with comprehension. For example, if children clearly cannot recognise common words in their language (as assessed by a word-list task), they will have trouble comprehending when it comes to reading a passage. If they cannot decode pseudo words, for example, they may have trouble processing an unfamiliar word in a passage even if they know the word orally, which in turn slows down comprehension. More complicated comprehension problems, such as the child not having an explicit strategy for comprehending, or not being accustomed to dialoguing around the meaning of text (or even the meaning of a passage read orally to them), are harder to assess even with oral assessments that have quite a few tasks. Overall, more effort needs to be dedicated to measuring reading comprehension in the early grades.

Finally, it is important to note that in many situations having additional reading comprehension questions may not be practical. UNICEF and Save the Children are exploring adding a few comprehension questions to the Multiple Indicator Cluster Survey (MICS), which would provide another measure of early grade reading among children aged 7-14 years in developing countries around the world (Cardoso and Dowd, 2016).

COMPARABILITY OF ORAL READING ASSESSMENTS

Although most oral reading assessments measure the same reading constructs, they are not



©USAID Lifelong Learning Project, Guatemala

necessarily comparable across countries and languages. In fact, differences in language structure and complexity make direct comparison of the results impractical, particularly direct comparisons of fluency. Comparing subtask results across countries and languages is therefore not advised—although it is possible to compare the percentages of children obtaining zero scores on specific tasks across languages and countries (Gove and Wetterberg, 2011). Although the inability to complete a task at *all* would not be affected by language structure and complexity, contextual factors such as exposure to print can lead to differences in zero scores.

Assessment results can, however, be used for *indirect* comparisons. For example, by comparing percentages of children reaching locally established benchmarks as opposed to percentages of children who reach a predetermined specific or international benchmark. The use of locally established benchmarks may provide countries and development partners with common ways to

measure and discuss progress towards the SDGs related to learning outcomes. Hsieh and Jeng (2016) explain how the government of The Gambia has been monitoring the progress of early grade reading using nationally-set benchmarks for reading in the language of instruction as well as other national languages. Benchmarks should be based on evidence from assessments that demonstrates that the levels of certain skills (or performance on certain metrics) are valid.

Even though most oral reading assessments are not designed to be comparable across countries or cultures, research shows that there are some skills and areas of development that can be compared. These ideas are being put into practice within the International Performance Indicators in Primary School (iPIPS) project, a cross-national oral reading assessment that captures non-cognitive development as well as cognitive skills (Merrell and Tymms, 2016).

RECOMMENDATION 4:

Properly organize the implementation of activities—logistics and monitoring

- Whether conducted in schools or households, common actions to maximise responses and assure data quality include: engaging a solid team to collect the data, providing adequate training and measuring inter-rater reliability aiming for minimum acceptable levels; providing clear instructions to assessors and properly documenting any actions taken; notifying target schools or villages/ households prior to the assessment; and gaining permission to assess the child ahead of time.
- Timely and consistent monitoring allows the teams to make adjustments during the fieldwork as it may be too late to fix issues that are identified only after the data collection has been completed.

Properly organized implementation of field activities cannot be underestimated. Organizing the implementation of the assessment includes logistics; monitoring the implementation and its progression; and steps to be taken after the assessment has been conducted.

CONDUCTING THE ASSESSMENT: LOGISTICS

Although logistics vary from country to country, there are some common practices that can maximise responses to the assessment, ensure a successful data collection and assure data quality. These include:

1. Engaging a knowledgeable team to collect the data

The individuals who conduct the field operations play a critical role and ultimately, the validity of the data will rely on the quality of their performance. These individuals generally include but are not limited to assessors, supervisors, scorers, data entry staff, drivers and other logistic support staff. They should be trained, have clearly defined roles and should not be expected to perform the impossible. The number of individuals performing the different tasks will vary on a case-by-case basis.

2. Notifying the schools or the villages/ households of the assessment

Implementers should contact the sampled schools to confirm their location, that they have pupils enrolled in the grade to be assessed and that the language

of instruction matches that of the assessment (Kochetkova and Dubeck, 2016). For assessments to be conducted in the home, common practices include announcing the household visit at the schools visited so that children are aware and wait for the team of assessors to arrive as well as asking the village leaders to inform households of the arrival of assessors prior to the visits (Mugo et al., 2016).

3. Taking into account the weather and terrain conditions

The weather conditions at the time of year that data collection will take place could impact fieldwork. For example, in some countries, the end of the school year may correspond with the rainy season or worsening road conditions and could potentially have an impact on school attendance or operation in certain areas (Kochetkova and Dubeck, 2016).

4. Ensuring clear instructions for actions to be taken when a sample school or child cannot engage in the assessment and carefully documenting and justifying all replacements

Replacement schools should be selected based on their similarity to the originally sampled school, such as location, type (public or private), enrolment, etc. Sampled schools that are located in difficult-to-reach areas should not be replaced simply for convenience—although in some cases, a sampled school will be totally unreachable by assessment teams due to weather or road conditions and will have to be replaced (Kochetkova and Dubeck, 2016).

TABLE 2

Possible solutions to common challenges encountered in household-based assessments

Challenge	Possible solutions
Parents are unhappy and reprimand children because they cannot read.	<ul style="list-style-type: none"> Assure parents that with time, children improve if they receive the required support. Work in pairs so one person can engage the parent in discussion away from the child while the other assesses the child.
Children fear reading in the presence of parents and other people; neighbours and passers-by disrupt the assessment.	<ul style="list-style-type: none"> Politely request those present to give assessors time with the child alone and tell them that the results will be explained after the assessment. Ask to take the child away from the crowd for testing with the permission of the child's relatives. Train volunteers to address disruptions.
Missing the assessment of many children because they cannot be found at home at the time of the visit.	<ul style="list-style-type: none"> Make callbacks later in the day or the following day. Announce household visits in the schools so that children are aware and wait for the team of assessors. Ask village leaders to inform households of the assessors' visits.
Households do not authorise assessing their children.	<ul style="list-style-type: none"> Take the time to introduce the assessment and the work, and seek consent. Use village leaders to inform the village prior to visits and if possible, also walk with the assessors during the assessment. Opt to use volunteers from the community.
Teachers are unaware of the learning issues captured by the assessment.	<ul style="list-style-type: none"> Share results during education days and visit some schools. Visit the government schools in the sampled villages to present the purpose of the survey and discuss the previous year's findings with teachers.

Source: adapted from (Mugo et al., 2016)

5. Gaining permission to assess the students/ children

For school assessments, gain the permission and trust of the school personnel, including the administration and the teachers. Negotiating a time to administer the assessments and respecting the wishes of the school is important. In many assessments, the explicit consent of each individual child is sought. In household based assessments, assessors must take the time to introduce the assessment to the household and seek consent from the parents and the children before proceeding (Mugo et al., 2016).

Finally, assessors administering a household-based assessment face a series of challenges during the data collection processes. **Table 2** offers practical solutions to common challenges faced during the administration of citizen-led assessments. This table is based on perspectives from the Annual Status of Education Report-India (ASER-India), ASER-Pakistan, Beekunko, Jàngandoo and Uwezo assessments.

QUALITY CONTROL AND MONITORING

Monitoring entails conducting a series of quality checks to ensure that the data collection is

progressing according to plan and that the assessors are administering the assessment in line with the guidelines provided. Monitoring assessor performance throughout the data collection process allows for timely intervention or retraining, which otherwise could go unnoticed until the end of the data collection process. Collecting data via electronic means can also help facilitate the early detection of problems.

In India, the ASER monitoring is done at two levels: one of the assessors by the supervisors or master trainers and the other of the master trainers by the state team (Banerji, 2016). It is also desirable—although not always possible—to measure inter-rater reliability during the fieldwork. Inter-rater reliability requires assessors to pair up to assess one of the selected children together each day. One interacts with the child while the other observes and marks the responses (Kochetkova and Dubeck, 2016). If this proves too expensive, inter-rater reliability can be assessed in training sessions where the trainer makes purposeful mistakes to see how the assessors perform and repeats the exercise until the rate of agreement among all assessors reaches a high percentage. The raters are considered reliable when the scores are the same or very close.

RECOMMENDATION 5:

Cater the analysis and communication of results to the target audience

- Report on the assessment results in a way that can be easily understood by a wide range of audiences, including non-experts. Descriptive statistics are a good technique for generating strong, easily grasped, readily communicable policy messages. However, results must be interpreted and reported with caution and should respect basic statistical principles.
- Determine the dissemination of activities at the country level based on two main factors: the purpose of the assessment and the audience. Consider a range of products to communicate the assessment results and use the appropriate language for dissemination, depending on the audience.
- Contributing to the knowledge base by sharing experiences through international platforms is encouraged; however, using the data generated to serve the country's own purposes and interventions is more important.
- Media campaigns are not to blame and shame. They should be used to publicise recommendations and strategies for how the system as well as children, parents and teachers can improve learning and disseminate key policy messages.

Analysing and interpreting the results is a crucial part of the assessment process. Additionally, presenting and communicating data from oral reading assessments to the right stakeholders is necessary to enable their use to inform decisions and design targeted interventions to improve reading. In a nutshell, analyses must be conducted and the results reported for different types of users—from policymakers to teachers looking to reinforce their pedagogical approaches and parents who want to work with their children to improve their learning. It is important to use the appropriate language for dissemination, depending on the audience.

ANALYSIS AND INTERPRETATION OF RESULTS

Measurement methods used to analyse and interpret the data depend on the test design of the assessment and the structure of the resulting data. In most oral reading assessments, performance results are analysed using descriptive statistics. This is possible since all children assessed are presented with an identical set of items. The use of descriptive statistics to analyse performance on an assessment typically entails reporting the percentage correct on a set of items. It therefore summarises the results in a format that is easily understood by a wide range of audiences and can be a good technique

for generating strong, easily communicable policy messages that can be readily grasped by other education stakeholders and non-experts. Common practices for analysing and reporting results include:

- Disaggregating results by specific characteristics of interest, such as grade, gender and geographical location. However, the number and levels of disaggregation are dependent on the data that has been collected.
- Comparing results against benchmarks that have been set either locally or internationally.
- Reporting results in a way that can easily be communicated and send a strong policy message. Some of the most common ways of reporting the data and the easiest to understand are:
 - mean scores by grade and subtask;
 - percentage of students in a given grade who can read the comprehension passage (within the allotted time, if timed) and correctly answer most or all reading comprehension questions;
 - percentage of students who scored zero by grade and region.
- Understanding how a range of variables interact. This has important policy consequences.

For example, if factors, such as low teacher absenteeism and principal management of student progress characterise good schools, then steps need to be taken to provide these opportunities to schools where the majority of pupils may be performing below the national benchmark. This may entail a combination of actions, such as providing resources and funding but also supporting schools that are accountable and well-managed.

- When possible, consider analysing the associations between different assessment results that are conducted on the same population of interest. For example, Banu Vagh (2016) examines the associations between ASER and EGRA conducted in India, to evaluate the validity of the ASER test. Although the two assessments are administered independently and have some differences, they are comparable in content as they are designed to assess the same abilities or skills.

It is important to remember to respect the basic principles of reporting data and making inferences about the results. For example, when comparing two groups, it is important to include significance tests along with the descriptive statistics (e.g. means and standard deviations) as this is needed to determine whether one can infer a statistically-significant difference between the two groups.

Although analysing data using descriptive statistics can yield valuable results, there are difficulties in comparing students from different grades (i.e. ceiling or floor effects), comparing student abilities over time, and describing the skills of students at specific ability levels. Thus, in some cases, it may be more appropriate to develop a literacy scale—one that describes skills on a continuum and can be used to compare children of different ages or grades or compare children’s abilities over time. However, to construct such a scale, a more complicated technique like item response theory (IRT) must be used. An example of using IRT to construct a literacy scale is detailed by Meiers and Mendelovits (2016).

COMMUNICATIONS MATERIALS AND DISSEMINATION STRATEGIES

Once the data have been analysed and interpreted, there is a range of products and formats in which the information could be communicated to the various stakeholders. These include but are not limited to mass media campaigns, policy briefs, data visualisations, short infographics, a national report, dissemination meetings, workshops, journal articles and conference presentations.

In general, the dissemination activities at the country level will be determined by two main factors: the purpose of the assessment and the audience. If, for example, the purpose of the assessment is to serve as a national or system-level diagnostic to design a policy reform, an intervention or a programme, then the audience of interest could be a ministry of education, donors, civil society, community leaders, academics, practitioners and teacher unions. Different activities can be undertaken with different groups, such as policy dialogue workshops, curriculum- or standard-review workshops, social mobilisation or mass media campaigns, project design workshops, policy briefs, press releases, journal articles and conference presentations. Even if the purpose of the assessment was to generate discussion at the national level and to spur ministries into action, the reporting of the results to schools and teachers can complement this promotion of awareness (RTI International, 2009). In India, the ASER Centre prepares a series of slides, presentations and notes for each state. State report cards are printed in a two- or four-page format for large scale distribution at different levels (Banerji, 2016). The media uses these documents to communicate the key findings by state to a wide audience.

INTERNATIONAL PLATFORMS

It is advised to make assessment reports publically available in order to help broaden the knowledge base of experiences in the development and application of oral reading assessments as well as their use. The following international platforms provide a wealth of information to practitioners,



© Margarita Montealegre, Nicaragua

international development agencies, governments, teachers associations, academics, civil society organizations, donor organizations, UN agencies, and other stakeholders:

- The United States Agency for International Development (USAID) EdData II website, developed to share experiences and reach a broad range of audiences at the international level.
- The UIS Catalogue and Database of Learning Assessments, compiles information on learning assessments which is organized in succinct formats to shed light on key characteristics of large-scale assessments and public examinations conducted in developing countries.
- The World Bank EdStats (Education Statistics) portal, a data and analysis source for key topics in education. It holds data from various sources which can be accessed through pre-defined data queries.
- The People's Action for Learning (PAL) Network bring together the countries working around the world to assess the basic reading and numeracy competencies of all children, in their homes, through annual citizen-led assessments. The PAL Network website provides relevant resources

from the citizen-led assessments, including research reports and assessment tools.

- The Global Reading Network brings together professionals from all over the world to improve children's reading. Resources and evidence-based practices for improving reading skills can be accessed on the Global Reading Network website.

THE MEDIA

How important is the media in disseminating the assessment results? Silvia Montoya, UIS Director, who formerly led several learning assessment initiatives in her native Argentina notes: 'media reports about learning assessment data make me cringe.' She explains in a blog post that the media should not be used to highlight 'bad grades', but rather to publicize recommendations and strategies for how children, parents and teachers can improve learning (Montoya, 2015). This type of media campaign was applied in Yemen and focused on encouraging parents and communities to support children's reading. Mobile service providers supported the campaign by broadcasting key messages to the 9 million subscribers (Creative Associates, 2016).

RECOMMENDATION 6:

Use the data to raise awareness and design interventions aimed at improving teaching and learning

- Stand-alone analyses may not be sufficient. Common practices to ensure that the data and the recommendations produced are considered for further action include: regular communication to promote community involvement, inviting change agents to participate in the data collection, encouraging local ownership, and consistent and regular assessment.
- In the process of improving reading skills, time, space and the mechanism must be given to stakeholders to engage with the assessment results. Change will not be immediate—improvements in reading achievements require time, resources and dedicated partners.
- Improving reading skills can be accomplished through the implementation of various programmes. Practices such as national campaigns as well as inciting parents, teachers and peers to improve teaching and learning have been successful. These practices can be applied in different contexts and settings.
- A programme design must be informed by the data.

Data from learning assessments can be packaged to serve various audiences, including policymakers and governments, teachers and parents. However, stand-alone analyses may not be sufficient.

PRODUCING LOCALLY OWNED DATA

Some common practices are encouraged to ensure that the data produced are owned by the community and the recommendations are considered for further action:

1. Regular communication with and involvement of the community

Communication should be regular with the different stakeholders throughout the different stages of the assessments. Parents, teachers, school personnel and ministry (government) officials need to be involved in the assessment strategy. The results of the assessment should be shared with parents, teachers and government officials. Teachers and parents should be provided with information on ways to improve reading skills.

2. Local ownership

Local ownership and participation is necessary to build awareness, improve accountability and initiate

action towards improving elementary education. Banerji (2016) stresses the importance of local ownership as an important element in the overall architecture of the ASER India. From inception, a key component of the ASER process was to involve local organizations and institutions. The local partners are involved in data collection as well as the dissemination of the results.

3. Consistent and regular assessment

Annual or regular cycles of assessment create a natural pulse of repetition where findings are regularly shared. This builds familiarity with the assessment among national policymakers, civil society organizations and advocacy groups and draws attention to the findings from year to year (Aslam et al., 2016).

IMPROVING READING SKILLS

Most of the organizations that participate in oral assessments emphasise that their purpose in assessing is to encourage the creation of programmes or interventions to improve reading skills and ensure that, through measurement, their chances of success are increased. Whether conducted in schools or households, data from oral reading assessments (i.e. learning achievement



data and the accompanying relevant contextual information) can and have been used to design strategies aimed at improving reading skills. The following are examples of good practices for using data from learning assessments to improve reading skills. These practices can be adapted to various settings, although feasibility and cost will differ depending on the context.

1. Reading campaigns

The collaboration among the different stakeholders that use the resulting assessment data to improve reading is just as important as the relationship between the partners during the planning and implementation of the assessment. The *Vamos a Leer, leer es divertido* ('Let's Read, reading is fun') is an ongoing campaign launched in 2010 in Nicaragua. It represents a collaborative effort between government, civil society, NGOs and private organizations to improve literacy in Grade 1 children. Their joint effort has created a culture of assessing reading to spark improvement and has

led to an increased number of libraries in schools (following the revelation that there was a shortage of reading materials in schools), helped teachers involve parents in reading and telling stories to their children, resulted in the development of several teacher training programmes, and, most importantly, demonstrated an improvement in early grade reading skills (Castro Cardenal, 2016).

2. Teachers

Although teachers could use the EGRA (or an adaptation of the assessment) and other multi-task assessments in their entirety, this is generally not recommended. More commonly, selected tasks are used as a type of formative assessment to monitor classroom progress, determine trends in performance and adapt instruction to meet children's instructional needs (Dubeck et al., 2016). Oral reading assessments designed to be conducted by teachers have had some positive effects on reading—they have provided important and useful insights into the progress and achievement of

students and has helped teachers adapt teaching/learning strategies to improve instruction. Rosales de Véliz et al. (2016) show how linguistic profiles have been developed and used in Guatemala to help teachers instruct reading in Spanish and other mother tongue languages. Meyer and Mendelovits (2016) show how longitudinal assessments that yield individually-reported results can provide teachers with a sound basis for planning future teaching strategies to meet the needs of their students.

Citizen-led initiatives—although conducted in households—have also involved teachers in the assessment process. These assessments have used teachers as assessors, enabling them to observe children’s weaknesses in the different learning processes and adopt counter measures in the classroom. In Senegal, in response to very low learning levels in Arabic-medium schools, Jángandoo has worked with school supervisors and the decentralised school authorities in one region to develop, test and introduce remedial education guides designed to provide teachers with new instructional approaches and materials (Ba et al., 2016).

3. Parents

Parents can play a valuable role in a literacy programme. Where possible, parents should receive training to help their children develop literacy skills. For example, the Yemen Early Grade Reading Approach (YEGRA) programme trained more than 23,000 parents on ways to support their children’s reading at home and to prepare children to attend school regularly and on time. The positive gains observed from the programme included improved reading skills in children (see **Box 5**). Parents reported that their children were reading or were being read to more at home and were more reluctant to miss school. The programme even influenced illiterate parents to learn how to read with their children (du Plessis et al., 2016). In addition, citizen-led assessments conducted in East Africa have helped shift the thinking of parents from assuming that learning is the sole responsibility of schools and teachers. It has helped raise awareness that parents have a major role to play in their children’s academic education. Citizen-led assessments have made a considerable effort to ensure that parents act on the advice provided by the assessors on

Box 5. Response to assessment results and findings

Programmes designed to improve reading skills generally involve various interventions, targeting different ‘change agents,’ including parents and teachers. The YEGRA programme was designed using the information from two oral reading assessments, EGRA and Literacy Boost.

EGRA and Literacy Boost assessment findings	YEGRA programme responses
Children who have regular attendance do better in reading.	The national media campaign and parent training messages included this statement: ‘Getting your children prepared for school in the morning and on time everyday helps student learning.’
Children who practice reading more, do better in reading.	All children have individual daily in-class reading.
Children who are read to at home or have books in the home perform better than those who don’t.	Training for parents in making the home a print rich environment, reading to children at home and ensuring they have opportunities to read outside the home (i.e. at mosques, libraries, shops and other places with public texts).
Regular corrective feedback to students is correlated with increased early grade reading scores.	Five assessments included in the teacher’s guide. One assessment administered after approximately every 20 lessons.
Student’s phonological awareness in Modern Standard Arabic is weak likely leading to poor uptake of letter sound recognition.	Teacher guides include focus on phonemic awareness with daily interactive practice for students.

Source: adapted from (du Plessis et al., 2016)



© Margarita Montealegre, Nicaragua

the importance of reading. For example, Uwezo assessors presented families with a calendar that included written suggestions of what parents can do to improve their children's learning (e.g. 'encourage your child to read at home') (Nakabugo, 2016).

4. Peers

Peers, especially those with more advanced skills, can be valuable allies in helping children develop literacy skills. The use of structured peer work has been researched extensively in developed countries (e.g. Dowhower, 1989; Fuchs et al., 1997) and has been included in reading programmes in developing countries. Peer work can be used to practice a number of literacy skills in the classroom. Outside of school, children in some cultures like to play school, which can help them learn. Although not specifically recommended by the authors, the role of peers is an avenue to explore and a good practice to consider for future interventions or programmes aimed at improving reading skills.

5. Reading materials

Data from oral reading assessments have also provided valuable information on developing reading materials to help improve reading skills. For example, the assessment in Guatemala has helped the development of various educational materials, such as *El tesoro de la lectura* ('The treasure in reading') series that addresses topics, such as emergent reading, reading development stages, associated skills and reading comprehension. The materials were further distributed to classrooms throughout the country and are expected to have a positive impact on learning to read (del Valle Catalán, 2016).



NUMERACY 26th FEBRUARY 2008
ADDITION WITH CARRYOVER

$$\begin{array}{r} 24 \\ + 14 \\ \hline 38 \end{array}$$

Conclusion

To achieve the SDG for education, governments will need more and better data to develop the evidence base needed to identify and effectively address weaknesses while monitoring progress. Early detection of learning gaps and support will be essential to inform remedial action and help realise the ambitious new global education goal of providing every child with a quality education and the foundational skills needed for a productive and fulfilling life. The key advantages that oral assessments provide are:

Timely access to data to inform decision making. Additional background information is often collected to support the results of the assessment and provide evidence to inform policy. The additional information collected depends on specific policy interests.

Early detection of reading weaknesses. Detecting learning problems early—particularly in reading as this will inevitably affect all other learning processes—allows for remedial action to be taken well before the end of primary education when it is often too late.

Viable solutions to measure the reading skills of children who are beginning to read. The assessments help capture the reading skills of children who have not yet mastered the necessary skills to take traditional written tests due to limited mechanical or decoding skills as well as comprehension and/or writing skills.

Means to assess learning for countries that do not participate in cross-national initiatives. The tools to conduct oral reading assessments are mostly open source, which allows practitioners to conduct an assessment at any time without

having to wait for the next cycle of cross-national assessments to become available (Gove, 2015). This also holds true for countries that do not participate in cross-national initiatives. However, oral reading assessments (unlike cross-national assessments) are not designed to be comparable across countries and especially across languages. This allows governments and their partners to conduct oral reading assessments at their discretion and without fear of being ranked or compared against other countries. Nevertheless, the accessible open-source availability of tools to conduct oral assessments does present a danger that an organization could apply the assessment carelessly and come to the wrong conclusions. The recommendations presented here have been produced to help address that concern. Along with the present document, many of the open sources for information on oral reading assessments contain detailed guidance that can help ensure that quality data are produced.

There is strong and systematic support from donors for countries measuring oral reading skills as a gateway to improved programmes and policies, stronger advocacy and better use of resources to improve learning outcomes. Further development of the generation and use of data from oral reading assessments must be encouraged through increased dialogue among implementers and practitioners. This will lead to a better understanding of what works and why, within and across countries.

References

The primary references of this report are the articles that are published in the ebook *Understanding What Works in Oral Reading Assessments*. The following is the list of articles and a short description of their content:

- 1. Aslam, M., Saeed, S., Scheid, P. and Schmidt, D. (2016). “Expanding citizen voice in education systems accountability: Evidence from the citizen-led learning assessments movement”.**
The article looks specifically at the ways in which citizen-led assessments have helped strengthen public accountability based on examples drawn from work in nine countries across South Asia, sub-Saharan Africa and Latin America. The article also describes the benefits citizen-led assessments afford education systems by generating nationally representative, locally generated and owned data on learning outcomes.
- 2. Ba, D., Bèye, M., Bousso, S., Mbodj, A. A., Sall, B. A. and Niang, D. (2016). “Evaluating reading skills in the household: Insights from the Jàngandoo Barometer”.**
This case study examines the Jàngandoo Barometer as a household-based citizen-led assessment administered in Senegal. It delves into the guiding principles of the assessment, the skills assessed, how the assessment informs interventions as well as the approaches used. The benefits and challenges associated with this assessment are presented as well as the remediation strategy initiative to address learning gaps.
- 3. Banerji, R. (2016). “Annual Status of Education Report (ASER) assessment in India: Fast, rigorous and frugal”.**
This case study discusses the planning, processes and the reasoning behind the structure of the ASER—a citizen-led, household-based assessment in India. The scope and scale of the ASER along with the challenges faced and impacts of the results are presented. The specific time frames used to conduct the ASER are also detailed.
- 4. Banu Vagh, S. (2016). “Is simple, quick and cost-also valid? Evaluating the ASER Hindi reading assessment in India”.**
This article evaluates the validity of the Annual Status of Education Report (ASER). Specific and detailed comparisons are made between the ASER Hindi reading assessment, the Early Grade Reading Assessment (EGRA) adapted for use in Hindi and the Read India (RI) pen-and-paper assessment of children’s basic and advanced Hindi literacy skills.
- 5. Beattie, K. and Hobbs, J. (2016). “Conducting an Early Grade Reading Assessment in a complex conflict environment: Is it worth it?”**
This article documents the experiences of Concern Worldwide in South-Central Somalia. The first Early Grade Reading Assessment (EGRA) in this conflict-ridden country was conducted in 2013 and provided baseline data to measure change from a literacy intervention. Results of the assessment are presented and the value of investing in literacy assessments in a country prone to conflict is discussed.

- 6. Cardoso, M. and Dowd, A.J. (2016). “Using Literacy Boost to inform a global, household-based measure of children’s reading skills”.**
This article explains the methodology to measure early reading skills among children using UNICEF’s Multiple Indicator Cluster Survey (MICS), including which constructs are targeted. It delves into a discussion on the number of comprehension questions to be used to measure each comprehension skill. The analytical methods presented focus on secondary analysis using data from Literacy Boost assessments conducted in seven developing countries.
- 7. Clark-Chiarelli, N. and Louge, N. (2016). “Teacher quality as a mediator of student achievement”.**
The article explores how to investigate teaching quality in primary grade literacy instruction. The article presents the Standards-based Classroom Observation Protocol for Educators in Literacy (SCOPE–Literacy), a tool that is proposed to help assess primary grade literacy instruction that goes beyond teachers’ credentials or content knowledge.
- 8. Castro Cardenal, V. (2016). “Use of literacy assessment results to improve reading comprehension in Nicaragua’s national reading campaign”.**
This article explores the impact of the national reading campaign in Nicaragua launched in 2010. The context, purpose and accomplishments of the reading campaign are discussed. The article delves into how the use of a version of the Early Grade Reading Assessment (EGRA) has helped create a culture of literacy assessments in Nicaragua.
- 9. del Valle Catalán, M. J. (2016). “Assessing reading in the early grades in Guatemala”.**
This case study details the reforms implemented following the results of a learning assessment in Guatemala. The article discusses how an oral reading assessment was adapted to test foundational reading skills and how the test results were used to design a learning intervention programme. The lessons learned from the experience as well as the next steps are also detailed.
- 10. Dowd, A. J. and Friedlander, E. W. (2016). “Home literacy environment data facilitate all children reading”.**
This article presents a field-tested method to collect learning environment data alongside oral reading assessment scores. The article underscores the importance of mapping children’s learning environments inside and outside schools to help build better learning interventions. The article also describes how to capture the learning environment in the context of the developing world and how to analyse the resulting data to inform efforts to improve learning.
- 11. Dowd, A. J., Pisani, L. and Borisova I. (2016). “Evaluating early learning from age 3 years to Grade 3”.**
This article explores how to assess children’s early learning status and their progress towards making meaning of text. It details two assessments—the Literacy Boost and the International Development and Early Learning Assessment. Together these assessments capture the continuum of language and literacy skill development from foundational to more advanced skills.
- 12. du Plessis, J., Tietjen, K. and El-Ashry, F. (2016). “The Yemen Early Grade Reading Approach: Striving for National Reform”.**
This case study delves into the role oral reading assessments have in sparking education reform in a country experiencing conflict. The authors also discuss four models of oral reading assessments used to design, evaluate and provide feedback on the Yemen Early Grade Reading Approach (YEGRA).

- 13. Dubeck, M. M., Gove, A. and Alexander, K. (2016). "School-based assessments: What and how to assess reading".**
The article looks at the reading skills that should be selected for inclusion in the Early Grade Reading Assessment (EGRA) and how the assessment generates results that are valid and reliable.
- 14. Hsieh, P. J. and Jeng, M. (2016). "Learning-by-doing: The Early Literacy in National Language Programme in The Gambia".**
This case study discusses the Early Literacy in National Language (ELINL) programme and its effect on the education system in The Gambia. It also delves into the different challenges faced in the development of the ELINL, including accounting for the diversified local languages used in a multilingual nation and producing comparable assessment results.
- 15. Kinyanjui, J. (2016). "Utility of the Early Grade Reading Assessment in Maa to monitor basic reading skills: A case study of Opportunity Schools in Kenya".**
To monitor pupils' progress in the acquisition of reading skills, the Women Educational Researchers of Kenya (WERK) developed an Early Grade Reading Assessment (EGRA) in the Maa language. This article presents how the assessment was adapted from the Research Triangle Institute (RTI) EGRA in English and the Uwezo framework for reading in Kiswahili. The successes and challenges faced from implementing the assessment are also detailed.
- 16. Kochetkova, E. and Dubeck, M. M. (2016). "Assessment in schools".**
This article provides an overview of the process of administering the Early Grade Reading Assessment (EGRA) in schools. It touches on the practical considerations when planning and conducting this type of assessment. The steps that are covered include recruiting and training assessors, performing a pilot test of the survey instruments, data collection and the protocols to follow when conducting school visits.
- 17. Meiers, M. and Mendelovits, J. (2016). "A longitudinal study of literacy development in the early years of school".**
This article discusses the Longitudinal Literacy and Numeracy Study (LLANS) undertaken by the Australian Council for Educational Research (ACER) from 1999 to 2005. It outlines the particularities of longitudinal studies and offers recommendations on planning and implementing such studies.
- 18. Merrell C. and Tymms P. (2016). "Assessing young children: Problems and solutions".**
This article details the experiences gained from over 20 years of developing the Performance Indicators in Primary Schools (PIPS), which has subsequently been expanded to an international project on the study of children starting school and the progress that they make during their first school year in different parts of the world. The article delves into the challenges of and possible solutions to assessing learning in young children, while fostering international comparisons when feasible.
- 19. Mugo, J. K., Kipruto, I. J., Nakhone, L. N. and Bobde, S. (2016). "Assessing children in the household: Experiences from five citizen-led assessments".**
This article explores the opportunities and challenges associated with assessing children at the household level through the experiences gained from five assessments: the ASER-India, ASER-Pakistan, Uwezo (Kenya, the United Republic of Tanzania and Uganda), Beekunko (Mali) and Jàngandoo (Senegal). The article describes a study that was conducted to gain insights from key informants who have experience implementing each of these five assessments.
- 20. Nakabugo, M. G. (2016). "What and how to assess reading using household-based, citizen-led assessments: Insights from the Uwezo annual learning assessment".**
This article outlines key features of household-based citizen-led assessments. It examines the

utility of this approach using insights gained from the Uwezo learning assessment—an East African initiative. The article details how and why household-based assessments measure basic reading competencies.

21. Rosales de Véliz, L., Morales Sierra, A. L., Perdomo, C. and Rubio, F. (2016). “USAID Lifelong Learning Project: The Linguistic Profile assessment”.

This article outlines the process of creating the Linguistic Profile assessment in Guatemala—a tool designed to help pre-primary and Grade 1 teachers determine the oral proficiency in Spanish and two Mayan languages of children entering the early grades in primary school. The article addresses how the results of the assessment can inform teachers’ decisions on the language of instruction and other bilingual education components in the first and second language.

22. Shirzad, H. and Magee, A. (2016). “Administering an EGRA in a post- and an on-going conflict Afghanistan: Challenges and opportunities”.

This case study explores the utility of the Early Grade Reading Assessment (EGRA) in Afghanistan. The strategic goals of Concern Worldwide are provided in a country-specific context along with the aims of the Rural Education Support Programme (RESP) launched in 2012. The challenges and advantages of administering the EGRA in a country with a history of conflict that is on-going are discussed.

ADDITIONAL REFERENCES

Fenton, R. (1996). “Performance assessment system development.” *Alaska Educational Research Journal*. Vol. 2, No. 1, pp. 13-22.

Chesterfield, R. and Abreu-Combs (2011). *Centers for Excellence in Teacher Training (CETT): Two-Year Impact Study Report (2008-2009)*. Washington DC: USAID Bureau for Latin America and the Caribbean. http://pdf.usaid.gov/pdf_docs/PDACS248.pdf

Creative Associates (2016). <http://www.creativeassociatesinternational.com/past-projects/yemen-early-grade-reading-approach/> (Accessed January 2016).

Dubeck, M. M. and Gove, A. (2014). *The early grade reading assessment (EGRA): Its theoretical foundation, purpose, and limitations*. Research Triangle Park, NC: RTI International. http://ac.els-cdn.com/S0738059314001126/1-s2.0-S0738059314001126-main.pdf?_tid=32ed8cfa-c9e3-11e5-a55a-0000aacb35d&acdnt=1454441515_8f5628dd667cfe473ba82558578c223d

Dowhower, S. L. (1989). “Repeated reading: Research into practice.” *The Reading Teacher*, Vol. 42, pp. 502-507.

Fuchs, D., Fuchs, L.S., Mathes, P.G. and Simmons, D.C. (1997). “Peer-Assisted Learning Strategies: Making classrooms more responsive to diversity.” *American Educational Research Journal*, Vol. 34, pp. 174-206.

Hart, B. and Risely, T.R. (2003). “The Early Catastrophe: The 30 Million Word Gap by Age 3.” *American Educator*, Spring 2003 pp. 4-9.

Global Reading Network. <https://www.globalreadingnetwork.net/>

Juel, C. (1988). “Learning to read and write: A longitudinal study of children in first and second grade.” *Journal of Educational Psychology*. Vol. 80, pp. 437-447.

Goswami, U. (2006). “Neuroscience and education: From research to practice.” *Nature Reviews Neuroscience*, Vol. 7, No. 5, pp. 406-413.

Gove, A. and Wetterberg, A. (eds.) (2011). *The Early Grade Reading Assessment: Applications and interventions to improve basic literacy*. Research Triangle Park, NC: RTI International.

- Hess, R. D. and Holloway, S. D. (1984). "Family and School as Educational Institutions." *Review of Child Development Research*, 7, 179–222.
- Montoya, S. (2015). *Why media reports about learning assessment data make me cringe*. Global Education Monitoring Report: World Education Blog. <https://efareport.wordpress.com/2015/06/17/why-media-reports-about-learning-assessment-data-make-me-cringe/>
- People's Action for Learning (PAL) Network. <http://palnetwork.org/>
- Sillers, D. (2015). USAID presentation at the 2015 Global Education Summit. <https://www.usaid.gov/sites/default/files/documents/1865/Sillers.pdf>. (Accessed January 2016).
- Wagner, D.A. (2011). "Smaller, Quicker, Cheaper: Improving Learning Assessments to Developing Countries." Paris: UNESCO-IIEP. <http://unesdoc.unesco.org/images/0021/002136/213663e.pdf>
- Wagner, D.A., Babson, A. and Murphy, K. M. (2011). "How Much is Learning Measurement Worth? Assessment Costs in Low-Income Countries." *Current Issues in Comparative Education*. Vol. 14: pp. 3-23.
- Research Triangle Institute International (2009). *Early Grade Reading Assessment toolkit*. USAID Education Data for Decision Making (EdData II). Washington, D.C.: USAID. <https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=149>
- Research Triangle Institute International (2014). *EdData II: Egypt Grade 3 Early Grade Reading—Pilot Group Assessment*. USAID Education Data for Decision Making (EdData II). Washington, D.C.: USAID. http://pdf.usaid.gov/pdf_docs/PA00K7GC.pdf
- UNESCO Institute for Statistics Catalogue of Learning Assessments. http://www.uis.unesco.org/nada/en/index.php/catalogue/learning_assessments. (Accessed January 2016).
- UNESCO Institute for Statistics Database. <http://www.uis.unesco.org/>. (Accessed January 2016).
- UNESCO TAG (2015). *Technical Advisory Group Proposal: Thematic Indicators to Monitor the Education 2030 Agenda*. Paris: UNESCO. <http://www.uis.unesco.org/Education/Documents/43-indicators-to-monitor-education2030.pdf>
- World Bank EdStats. <http://datatopics.worldbank.org/education/>
- Yoon, K. S., Duncan, T., Lee, S. W. –Y., Scarloss, B., and Shapely, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. Washington, DC: U.S. Department of Education, Institute of Educational Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf



Quality education and learning for all is at the center of the Sustainable Development Goal (SDG) for education. Policymakers at the global and national levels clearly recognise that to determine if the quality of education is improving, they must first engage in learning assessments to monitor learning outcomes. Oral reading assessments play an important role in measuring the basic foundational reading skills that are a gateway skill to all other learning.

Today, the use of oral assessments is widespread and while there are some commonalities among the instruments used, there are also differences in purpose, design and administration. In response, the UNESCO Institute for Statistics (UIS) led a collaborative effort with organizations that have been actively financing, designing and implementing oral assessments. Representatives from the participating organizations submitted case studies and position papers to help exemplify good practices in oral reading assessments.

This report presents their recommendations for selecting, implementing and using learning assessments as well as basic principles that should be applied at the different stages of oral reading assessments—from planning and design to implementation and use of the resulting data.

As the SDGs become a reality, governments will need more and better data on education to produce evidence, determine the areas of improvement, take corrective action and monitor progress. Early detection of learning gaps will be essential to informing remedial action and securing the ambition of the new goals to ensure that all children gain access to post-primary education. This report serves as a unified voice from the community of oral reading assessment practitioners, implementers and donors on the importance of early reading skills to ensure learning for all by 2030.

To access the ebook featuring the collection of articles on oral reading assessments, please visit www.uis.unesco.org

