Toward Universal Learning

A Global Framework for Measuring Learning

2

2

Toward Universal Learning

A Global Framework for Measuring Learning

UNESCO

UNESCO INSTITUTE *for* STATISTICS

United Nations
Educational, Scientific and
Cultural Organization

B | Center for
Universal Education
at BROOKINGS

Toward Universal Learning: A Global Framework for Measuring Learning is the second in a series of three reports from the Learning Metrics Task Force. The first report was focused on identifying what learning is important for all children and youth (*Toward Universal Learning: What Every Child Should Learn*), and a third report will address how the measurement of learning can be implemented to improve education quality.

This report represents the collaborative work of the Learning Metrics Task Force's members and their organizations, a technical working group convened by the task force's Secretariat, and more than 600 individuals around the world who provided feedback on the recommendations. The members of the Measures and Methods Working Group who wrote the report are listed below.

About the Learning Metrics Task Force

The UNESCO Institute for Statistics and the Center for Universal Education at the Brookings Institution have joined efforts to convene the Learning Metrics Task Force. The overarching objective of the project is to catalyze a shift in the global conversation on education from a focus on access to access *plus* learning. Based on recommendations from technical working groups and input from broad global consultations, the task force works to ensure that learning becomes a central component of the post-2015 global development agenda and to make recommendations for common goals to improve learning opportunities and outcomes for children and youth worldwide. Visit www.brookings.edu/learningmetrics to learn more.

This is a joint publication of the UNESCO Institute for Statistics and the Center for Universal Education at the Brookings Institution. Any citation of this report should include specific reference to both organizations. The following is a suggested citation:

The 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 program 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.

The Center for Universal Education at the Brookings Institution

The Center for Universal Education (CUE) at the Brookings Institution is one of the leading policy centers focused on universal quality education in the developing world. CUE develops and disseminates effective solutions to achieve equitable learning, and plays a critical role in influencing the development of new international education policies and in transforming them into actionable strategies for governments, civil society and private enterprise. The Center for Universal Education is engaged in four broad areas: influencing the global education to 2015 and beyond; improving education resources and learning outcomes; advancing quality education for the marginalized; and promoting collaboration between diverse stakeholders in the education sector.

The Brookings Institution is a private, nonprofit organization. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars. Brookings recognizes that the value it provides is in its absolute commitment to quality, independence and impact. The activities supported by its donors reflect this commitment, and the analysis and recommendations are not determined or influenced by any donation.

The Learning Metrics Task Force

Cochairs	
Rukmini Banerji, Director of Programs	Pratham
Sir Michael Barber, Chief Education Advisor	Pearson
Geeta Rao Gupta, Deputy Executive Director	UNICEF

Member Organizations and Representatives*	
ActionAid and GPE Board Representative for Northern Civil Society	David Archer, Head of Programme Development
African Union (AU)	Beatrice Njenga, Head of Education Division
Agence Française de Développement (AFD)	Valérie Tehio, Project Manager, Education
Arab League of Educational, Cultural, and Scientific Organization (ALECSO)	Abdullah Hamad Muhareb, Director General
Association for Education Development in Africa (ADEA)	Dzingai Mutumbuka, Chair
Australian Agency for International Development (AusAID)	Raymond Adams, Special Advisor, ACER; Professor, University of Melbourne
Campaign for Female Education in Zambia (Camfed) and Representative for Southern Civil Society	Lucy Lake, Chief Executive Officer
City of Buenos Aires, Argentina	Mercedes Miguel, General Director of Education Planning
Dubai Cares / United Arab Emirates (UAE)	H.E. Reem Al-Hashimy, Chair and Minister of State
Education International (EI)	Fred van Leeuwen, General Secretary
Global Partnership for Education (GPE)	Carol Bellamy, Chair of the Board
Government of Assam, India	Dhir Jhingran, Principal Secretary
Inter-American Development Bank (IADB)	Emiliana Vegas, Chief, Education Division
International Education Funders Group (IEFG)	Chloe O'Gara, Co-chair
Kenyan Ministry of Education	George Godia, Permanent Secretary
Korea Institute for Curriculum and Evaluation (KICE)	Seong Taeje, President
Office of the UN Secretary General	Itai Madamombe, Global Education Advisor
Organización de Estados Iberoamericanos (OEI)	Álvaro Marchesi, Secretary-General
Queen Rania Teacher Academy	Tayseer Al Noaimi, President and former Minister of Education of Jordan
Southeast Asian Minister of Education Organization (SEAMEO)	Witaya Jeradechakul, Director
UK Department for International Development (DFID)	Chris Berry, Senior Education Advisor
UNICEF	Jo Bourne, Associate Director, Education
United Nations Development Program (UNDP)	Selim Jahan, Director of Poverty Practice
UNESCO	Maki Hayashikawa, Chief, Section for Basic Education, Division for Basic Learning and Skills Development
United States Agency for International Development (USAID)	John Comings, Education Advisor
World Bank	Elizabeth King, Director of Education

as of April 2013

Secretariat	
UNESCO Institute for Statistics	
Hendrik van der Pol	Director
Albert Motivans	Head of Education Indicators and Data Analysis Section
Maya Prince	Research Assistant

Center for Universal Education at the Brookings Institution	
Rebecca Winthrop	Senior Fellow and Director
Xanthe Ackerman	Associate Director
Kate Anderson Simons	Technical Lead, Learning Metrics Task Force
Mari Soliván	Project Manager, Learning Metrics Task Force
Khaled Fayyad	Project Assistant
Lauren Lichtman	Research Assistant, Learning Metrics Task Force

Measures and Methods Working Group Members

Measures and Methods Working Group Chair: César Guadalupe, Universidad del Pacífico (Peru)

Secretariat Technical Leads: Kate Anderson Simons, Center for Universal Education at the Brookings Institution; Maya Prince, UNESCO Institute for Statistics

Early Childhood Subgroup Members	
Abbie Raikes	UNESCO, France
Aglaia Zafeirakou	GPE Secretariat, United States
Daniel Ansari	Western University, United Kingdom
Jennifer Blinkhorn	Agha Khan Foundation, Afghanistan
Magdalena Janus	Offord Centre for Child Studies, McMaster University, Canada
Mariana Hi Fong de Camchong	Universidad Casa Grande/Blossom Centro Familiar, Ecuador
Myung-Lim Chang	Korea Institute of Child Care and Education, South Korea
Nitika Tolani-Brown	Save the Children, United States
Nurper Ulkuer	Consultant, Turkey
Paula Louzano	Education Consultant, Brazil
Pierre Petignat	Université HEP-BEJUNE, Switzerland
Santiago Cueto	GRADE, Peru
Venita Kaul	Ambedkar University Delhi, India

Primary Subgroup Members	
Aanchal Chomal	Azim Premji Foundation, India
Amy Jo Dowd	Save the Children, United States
Antoine Marivin	PASEC, Senegal
Audrie-Marie Moore	FHI 360, United States
Benjamin Piper	RTI International, Kenya
Chong Min Kim	Korean Educational Development Institute, South Korea
Diana Weber	SIL International, United States
Edward Amankwah	Youth Action International, Ghana
Galina Kovaleva	Russian Academy of Education, Russia
Helen Abadzi	Global Partnership for Education, United States
Ina V.S. Mullis	TIMSS and PIRLS International Study Center, Boston College, United States
Joyce Kinyanjui	Women Educational Researchers of Kenya: Opportunity Schools, Kenya
Juan E. Jiménez	University of La Laguna, Spain
Kenneth Ross	University of Melbourne, Australia
Luis Crouch	Global Partnership for Education, United States
Margarita Peña Borrero	Instituto Colombiano para la Evaluación de la Educación, Colombia
Marguerite Clarke	World Bank, United States
Mary Anne Drinkwater	OISE, University of Toronto, Canada
Nadir Altinok	University of Lorraine, France
Roland Tormey	UNECE, Switzerland
Ruth Naylor	CfBT Education Trust, UK
Silvia Montoya	Buenos Aires City Government, Argentina
Vanessa Sy	PASEC, Senegal
Yanhong Zhang	UNESCO, Bangkok

Postprimary Subgroup Members	
Abdullah Ferdous	American Institutes for Research, United States
Bill Oliver	Voluntary Service Overseas, Papua New Guinea
Boban Varghese Paul	Pratham, India
Brenda S. H. Tay-Lim	UNESCO Institute for Statistics, Canada
Chang Hwan Kim	Korean Educational Development Institute, South Korea
Cynthia Lloyd	Population Council, United States
Diederik Schönau	Cito Institute for Educational Measurement, The Netherlands
Kateryna Shalayeva	International Development and European Law Consultant, Ukraine/France
Margaret Sinclair	Education Above All, Qatar
Marina L. Anselme	The RET, Switzerland
Markus Broer	American Institutes for Research, United States
Michael O. Martin	TIMSS and PIRLS International Study Center, Boston College, United States
Mi-Young Song	Korea Institute for Curriculum and Evaluation, South Korea
Pauline Greaves	Commonwealth Secretariat
Ralph St. Clair	McGill University, Canada
Ruth Mbabazi Kabutembe	Voluntary Service Overseas, Rwanda
Sarah Howie	University of Pretoria, South Africa
Solomon Nuni Esokomi	Teacher Service Commission, Kenya
Trey Menefee	Hong Kong University, Hong Kong
Usman Ali	Idara-e-Taleem-o-Aagahi (ITA), Pakistan
William G. Brozo	George Mason University, United States

Acknowledgements

Abbreviations and Acronyms

ASER	Annual Status of Education Report
CONFEMEN	Conférence des ministres de l'Éducation des pays ayant le français en partage
CoP	community of practice
CUE	Center for Universal Education
EDI	early development instrument
EFA	Education for All
EGMA	Early Grade Math Assessment
EGRA	Early Grade Reading Assessment
GER	gross enrollment ratio
GPE	Global Partnership for Education
ICCS	International Civic and Citizenship Education Study
IEA	International Association for the Evaluation of Educational Achievement
INEE	Inter-Agency Network for Education in Emergencies
ISCED	International Standard Classification of Education
LAMP	Literacy Assessment and Monitoring Programme
LLECE	Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación
LMTF	Learning Metrics Task Force
MDG	Millennium Development Goal
MICS	Multiple Indicator Cluster Survey
NAC	National Assessment Center
NER	net enrollment rate
NIR	net intake rate
OECD	Organization for Economic Cooperation and Development
OLA	Out-of-School Youth Literacy Assessment
PASEC	Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN
PIAAC	Program for the International Assessment of Adult Competencies
PIRLS	Progress in International Reading Literacy Study
PISA	Program for International Student Assessment
SABER	Systems Approach for Better Educational Results
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SDSN	Sustainable Development Solutions Network
TIMSS	Trends in International Mathematics and Science Study
UIS	UNESCO Institute for Statistics
UNESCO	United Nations Educational Cultural and Scientific Organization
UNICEF	United Nations Children's Fund

Contents

Introduction

The benefits of education—for national development, individual prosperity, health and social stability—are well known, but for these benefits to accrue, children in school need to be learning. Despite commitments and progress in improving access to education at the global level, including Millennium Development Goal (MDG) 2 on the universal completion of primary education and the Education for All (EFA) Goals, levels of learning are still too low. According to estimations in the 2012 EFA Global Monitoring Report, at least 250 million primary-school-age children around the world are not able to read, write or count well enough to meet minimum learning standards, including those who have spent at least four years in school (UNESCO 2012). Worse still, this figure is likely to be an underestimate because measurement of learning outcomes among children and youth is limited and, relative to the measurement of access, more difficult to assess at the global level.

To advance progress for children and youth around the world, it is critical that learning is recognized as essential for human development. As EFA and the MDGs sunset in 2015, and the UN Secretary-General promotes the Global Education First Initiative, the education sector has an opportunity to raise the profile of international education goals and ensure that learning becomes a central component of the global development agenda. One of the first major milestones in this process was the May 2013 release of the UN Secretary-General's High-Level Panel on the Post-2015 Development Agenda. The report is one part of a longer process for deciding the next global development agenda, but it presents a clear focus on education, and reflects the desire to measure progress using global measurements of access and learning outcomes at the primary and secondary levels (Secretary-General's High-Level Panel of Eminent Persons, SG HLPEP, 2013). It provides a framework for bringing together the international community's human development and sustainable development efforts, which to date have largely been separate. This report proposes illustrative global education goals and accompanying targets, including:

1. Increase by x% the proportion of children able to access and complete preprimary education;

2. Ensure that every child, regardless of circumstance, completes primary education able to read, write and count well enough to meet minimum learning standards;

3. Ensure that every child, regardless of circumstance, has access to lower secondary education and increase the proportion of adolescents who achieve recognized and measurable learning outcomes; and

4. Increase the number of young and adult women and men with the skills, including technical and vocational skills, needed for work by x% (SG HLPEP 2013, 36).

To ensure that these learning targets are measurable, the global education community must work collectively to define global ambition on improving learning and propose practical actions to deliver and measure progress. In response to this need, 30 organizations and more than 1,000 individuals around the world have come together to form the Learning Metrics Task Force, co-convened by UNESCO, through its Institute for Statistics (UIS), and the Center for Universal Education (CUE) at the Brookings Institution. The overarching objective of the project is to catalyze a shift in the global conversation on education from a focus on access to access *plus* learning. Based on recommendations of technical working groups and input from broad consultations, the task force aims to make recommendations to help countries and international organizations measure and improve learning outcomes for children and youth worldwide.

The task force—comprised of representatives of national and regional governments, EFA-convening agencies, regional political bodies, civil society organizations, and donor agencies—is engaged in an 18-month-long global consultation process to build consensus around three essential questions addressed in the following order:

- Phase I: What learning is important for all children and youth?

- Phase II: How should learning outcomes be measured?

- Phase III: How can the measurement of learning improve education quality?

This report is structured as follows. It first provides an outline of the purpose of the LMTF and an overview of the framework of the seven domains of learning. Then, this report describes the six areas of measurement—including the rationale and feasibility—and highlights some of the existing measurement efforts in each of the areas. Next, it discusses considerations for tracking learning goals in a way that focuses on equity. The subsequent section describes considerations related to the methodologies for learning assessments. The following section discusses the relationship between the six areas of global measurement and the current post-2015 global education and development dialogues. Finally, the report presents overarching considerations for the final working group on implementation.

Building Consensus on Measurement at the Global Level

The Education for All (EFA) goals, initiated in 1990 in Jomtien, Thailand, demonstrated a commitment to meeting basic learning needs. This commitment was restated in 2000 in the Dakar Framework for Action, in which Goal 6 states: "Improving every aspect of the quality of education, and ensuring their excellence so that recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy and essential life skills."

Measuring can play a crucial role in improving education quality and learning. Good teachers measure learning in the classroom to adjust and individualize instruction. Good head teachers, school administrators and school district leaders measure learning at the school and community levels to target resources and improve school quality. Many national governments measure learning to diagnose the overall health of the national education system and develop policies to improve learning outcomes. Civil society actors, donors, and development agencies use assessments to measure the effectiveness of programming and advocate for effective education policies and practices.
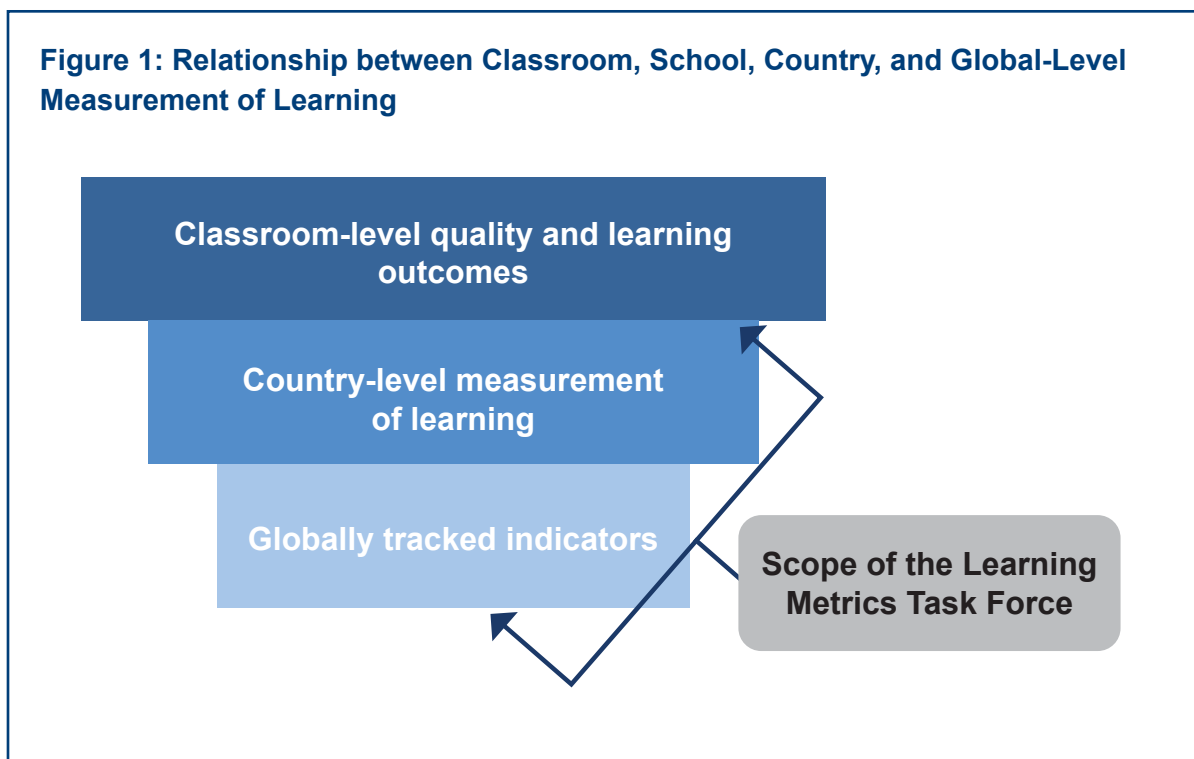
Nevertheless, the connection between measurement and the improvement of learning is neither automatic nor simple; different approaches to measurement are useful for different purposes. For measurement to be effective, it must be fit for the purpose. Therefore, no single approach to measuring learning is better than another; rather, the choices about what and how to measure should be informed by the needs of those who require the information. While measurement may have different purposes at different levels, the systems for measuring and improving learning at the classroom, national, and global levels should not be working in isolation. Globally tracked indicators should be aligned with what is measured nationally and in schools or classrooms, while measurement at the national level should be aligned with the competencies measured in classrooms or schools. As a global task force, the LMTF seeks to make recommendations at the national and global levels while acknowledging that the types of learning measured at the subnational and local levels are critical yet beyond the scope of the LMTF. Figure 1 represents the way these different levels coexist and the focus of the LMTF.

Process of the LMTF

In Phase I of the project, the LMTF's Standards Working Group convened from May to October 2012 to make recommendations on what learning competencies are important for children and youth around the world. The Standards Working Group circulated prototype recommendations for public consultation from August through September 2012 and modified them based on feedback from more than 500 individuals in 57 countries. A draft framework was presented to the task force at a two day in-person meeting in September 2012. Over these two days, the LMTF finalized a framework to be used by the subsequent working group on measures and methods to investigate the measurement of learning outcomes. The Standards Working Group was tasked with developing a framework for learning outcomes that would not be restricted to those outcomes that lend themselves easily to measurement and are, as a result, currently prioritized.

Figure 1: Relationship between Classroom, School, Country, and Global-Level Measurement of Learning

Classroom-level quality and learning outcomes

Country-level measurement of learning

Globally tracked indicators

Scope of the Learning Metrics Task Force

In Phase II of the project, the task force considered approaches for measuring and tracking progress in learning at the global and national levels—drawing on the framework elaborated by the first working group. The Measures and Methods Working Group—comprised of 57 experts in education, learning assessment and other relevant fields—provided technical guidance and recommendations for the task force to consider.

On February 20-21, 2013, the working group presented its recommendations on this topic to the task force at a meeting in Dubai hosted by Dubai Cares. Among the 44 attendees were representatives of low-, middle- and high-income countries; stakeholders from Africa, Asia, Europe, the Middle East, North and South America,

and Oceania; key UN and multilateral agencies; regional bodies; teacher organizations; civil society organizations; and bilateral donor agencies. In addition, more than 600 individuals in 57 countries around the world submitted comments on the initial draft of recommendations through a wide-ranging public consultation process. The third and final phase of the LMTF began in March 2013 with the launch of the Implementation Working Group. This working group presented its recommendations to the task force at its July 2013 meeting at the Rockefeller Center in Bellagio, Italy, and will finish its report by November 2013. Figure 2 describes the global reach of the working group members and consultation participants.

Figure 2: Global Reach of LMTF Consultation

Phase I: Standards

- Working Group: 39 members in 23 countries
- Consultation: 500 people in 57 countries, 75% Global South

Phase II: Measures and Methods

- Working Group: 57 members from 27 countries
- Consultation: 600 people from 57 countries, 50% Global South

Phase III: Implementation

- Working Group: 125 members in 40 countries
- Consultation: Nearly 700 people from 72 countries, 80% Global South

Toward Universal Learning: Measuring to Improve Learning is the second in a series of three reports by the Learning Metrics Task Force. The first report, *Toward Universal Learning: What Every Child Should Learn*, focuses on identifying what particular types of learning are important for every child (LMTF 2013). The p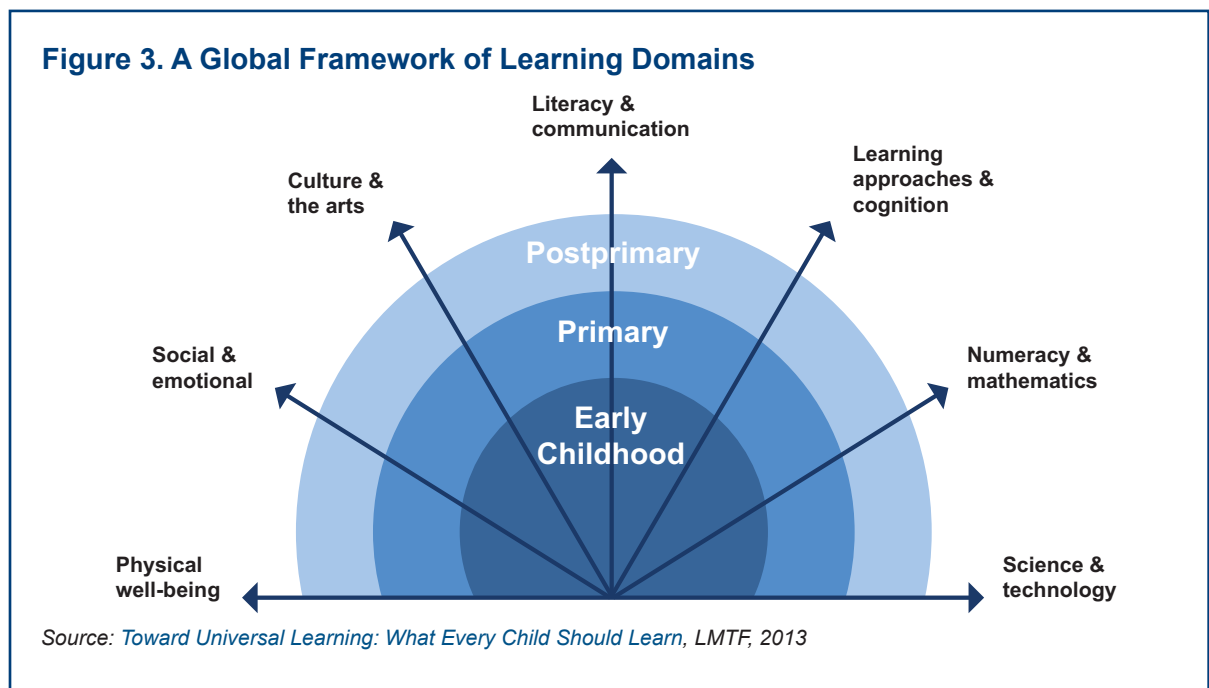resent report intends to provide guidance on how learning outcomes should be measured and builds upon the domains of learning identified in the first phase of the LMTF. The final phase will yield the third report in this series, and will address implementation-related issues to ensure that measurement informs action to improve learning.

What Learning Is Important for All Children and Youth?

Phase I of the project sought to answer the question, What do all children and youth need to learn in order to succeed in the 21st century? Considering recommendations from a working group of experts, the task force decided at its first meeting in September 2012 that, indeed, there were important competencies that all children and youth should master no matter where they live in the world. The first report from the task force, *Toward Universal Learning: What Every Child Should Learn*, presents a broad, holistic framework of seven learning domains as the aspiration for all children and youth around the world (LMTF 2013).

Given the diversity of structures, places and times in which children and youth learn, it is a challenge to de-

fine what outcomes related to learning are important, especially at the global level. Furthermore, in order to develop a framework that can stay relevant for the next 15 years, the task force recognized that it would need to take a step back from what is measurable today and first consider what learning is important for the 21st century. The feedback from interviews with key stakeholders and global consultations points to a growing demand for measuring learning in multiple areas and in different ways, not only limited to standardized tests of literacy and numeracy. Accordingly, the task force proposes a broad definition of learning that encompasses seven domains of outcomes, with corresponding subdomains, as important for all children and youth to develop (see figure 3).

Figure 3. A Global Framework of Learning Domains

Literacy & communication

Culture & the arts

Learning approaches & cognition

Postprimary

Primary

Social & emotional

Numeracy & mathematics

Early Childhood

Physical well-being

Science & technology

Source: *Toward Universal Learning: What Every Child Should Learn*, LMTF, 2013

This holistic framework of learning domains was developed based on:

- Existing global policies and dialogues, such as EFA and the UN Convention on the Rights of the Child, which mandate a broad definition of education and learning.

- Research supporting the importance of learning in these domains for different areas of people's lives, including economic growth and material prosperity.

- Results from global public consultation, in which more than 500 individuals in 57 countries provided feedback. The overwhelming majority of participants in the global consultation, especially those from the Global South, argued for a broad definition of learning that goes beyond basic literacy and numeracy.

The seven domains and corresponding subdomains are detailed in table 1. A full description of the seven domains and the methodology behind the framework is found in *Toward Universal Learning: What Every Child Should Learn* (LMTF 2013).

The task force decided that the subsequent working group should investigate ways to measure learning in all seven domains and at the three stages (early childhood, primary and lower secondary) and make recommendations on the feasibility for measuring learning at the global level. Lower secondary was selected as the upper end of the task force's recommendations due to the complex areas of specialization that occur beyond that level.

The public consultations demonstrated a need for common terminology when discussing learning metrics. For the purposes of this report, the LMTF uses the following definitions of key terms, each shown in **bold type**, related to measuring learning:

Goals are aspirational statements that can be used to motivate or inspire action. An example of a goal is, "All children start primary school ready to learn across five domains." Different tools, or **measures**, are administered to gauge progress toward the goal, and taken together the data from these measures provide a **metric** for the whole population. For example, a national assessment of reading is a measure used by countries to quantify children's knowledge and competencies in reading, and the results are combined to provide a national metric on reading outcomes. Within these measures, different levels of learning are possible. Some assessments, such as many national exams, have only two levels (e.g., pass or fail), while others have progressively complex levels of learning, called **benchmarks**. These benchmarks correspond to various levels of proficiency, such as basic, proficient, or advanced. An example of a possible basic benchmark is "Children can add single digit numbers." A higher benchmark on the same scale might be, "Children can multiply single digit numbers." Finally, **targets** are used to track progress toward the goals. An example of a target is, "By 2020, decrease by half the number of children unable to demonstrate reading competencies at an intermediate benchmark."

Table 1. Domains and Subdomains of Learning by Educational Stage			
Domain	**Subdomains**		
	Early Childhood	**Primary**	**Postprimary**
Physical well-being	• Physical health and nutrition • Health knowledge and practice • Safety knowledge and practice • Gross, fine, and perceptual motor	• Physical health and hygiene • Food and nutrition • Physical activity • Sexual health	• Health and hygiene • Sexual and reproductive health • Illness and disease prevention
Social and emotional	• Self-regulation • Emotional awareness • Self-concept and self-efficacy • Empathy • Social relationships and behaviors • Conflict resolution • Moral values	• Social and community values • Civic values • Mental health and well-being	• Social awareness • Leadership • Civic engagement • Positive view of self and others • Resilience/"grit" • Moral and ethical values • Social sciences
Culture and the arts	• Creative arts • Self- and community-identity • Awareness of and respect for diversity	• Creative arts • Cultural knowledge	• Creative arts • Cultural studies
Literacy and communication	• Receptive language • Expressive language • Vocabulary • Print awareness	• Oral fluency • Oral comprehension • Reading fluency • Reading comprehension • Receptive vocabulary • Expressive vocabulary • Written expression/ composition	• Speaking and listening • Writing • Reading
Learning approaches and cognition	• Curiosity and engagement • Persistence and attention • Autonomy and initiative • Cooperation • Creativity • Reasoning and problem solving • Early critical thinking skills • Symbolic representation	• Persistence and attention • Cooperation • Autonomy • Knowledge • Comprehension • Application • Critical thinking	• Collaboration • Self-direction • Learning orientation • Persistence • Problem Solving • Critical decisionmaking • Flexibility • Creativity

Domain	Subdomains		
	Early Childhood Level	**Primary Level**	**Postprimary Level**
Numeracy and mathematics	• Number sense and operations • Spatial sense and geometry • Patterns and classification • Measurement and comparison	• Number concepts and operations • Geometry and patterns • Mathematics application	• Number • Algebra • Geometry • Everyday calculations • Personal finance • Informed consumer • Data and statistics
Science and technology	• Inquiry skills • Awareness of the natural and physical world • Technology awareness	• Scientific inquiry • Life science • Physical science • Earth science • Awareness and use of digital technology	• Biology • Chemistry • Physics • Earth science • Scientific approaches • Environmental awareness • Digital learning

A Global Vision for Measuring Learning

After Phase I, the task force concluded that it is important to reinforce the importance of all seven domains and to emphasize that the comprehensive nature of learning should not be oversimplified by focusing on some domains over others. As a result, the task force charged the Measures and Methods Working Group with proposing a hybrid model for measuring learning, with some domains to be measured at the global level and others to be measured at the national level. While the curricula in many countries encompasses these domains (even though different nomenclature or categorization may be used), the working group acknowledged that measurement is currently not equally developed in each of the seven domains. As a result, the task force emphasizes the importance of all seven domains in promoting the holistic development of each person and that measurement efforts should ideally address all domains. The current feasibility of measuring these domains and subdomains is considered as one of the criteria pertaining to the areas of global measurement described in this report.

Nevertheless, the LMTF acknowledges that the different domains require varying degrees of effort to develop suitable measures at the national and international levels. Additionally, some domains might only be emphasized within specific national policy strategies in light of particular national conditions. For example, a country with high rates of HIV/AIDS may prioritize competencies related to physical health and well-being, and a country recently emerging from conflict may prioritize competencies in the social and emotional domain.

Based on the availability of national and international measures and feedback from the consultation process, the working group recommended to the task force 10 indicators that were feasible and/or desirable for tracking at the global level (see annex B). The 10 proposed indicators are in relation to the following competencies or domains:

- *Early childhood*: (1) school readiness (across five domains: physical well-being, social and emotional, literacy and communication, numeracy and mathematics, learning approaches and cognition)

- *Primary*: (2) social and emotional competencies, (3) reading, (4) mathematics, (5) science

- *Lower secondary*: (6) citizenship, (7) reading, (8) collaborative problem solving, (9) mathematics, (10) science

The task force concluded that deciding on specific indicators was not possible at the Dubai meeting but instead agreed upon six areas of measurement to track at the global level, which are described in detail in this report. The LMTF will continue to develop and refine the indicators for these six areas of measurement.

The working group also proposed a global mechanism, such as a multi-stakeholder advisory group, that could support countries in measuring learning according to national priorities. The working group chair and facilitators presented these recommendations to the task force at the February meeting in Dubai (see annex D for the preliminary recommendations from the working group).

The task force came to a consensus on the following decisions at the Dubai meeting, which are described in detail in the next section this report:

Six areas of measurement should be tracked at the global level.

After reviewing the 10 indicators proposed by the working group listed in the previous section (see annex B), the task force deliberated on which areas of measurement were feasible and desirable to be tracked at the global level, acknowledging that the entire framework of learning domains developed in Phase I is too large and complex for this purpose. Of the areas originally proposed, the six areas of measurement represent important learning opportunities for children and youth to succeed in a globalized society. The first two areas capture contexts for the learning process, and the last four describe demonstrable learning outcomes. The current feasibility of measurement in these areas is varied, with indicators within access and completion being almost universally tracked and other areas such as breadth of learning opportunities that are less developed and not currently tracked at the global level. These six areas represent the task force's vision for how learning should be measured globally, realizing the necessity for significant improvements in assessment capacity in many countries before all areas could be measured.

1) Access to and completion of learning opportunities.

Rationale: Tracking progress in access and completion of learning opportunities addresses the unfinished access agenda for out-of-school children and youth. It also allows for a broad definition of schooling, including criteria based on intentional learning programs, whether formal or nonformal. These programs occur

when individuals intentionally or voluntarily search for knowledge, skills, competencies or attitudes of lasting value, and that intention is formulated by the learner before starting the activity (European Commission 2006a). Evidence shows that the skills and knowledge needed to participate in a global economy are rarely acquired outside intentional learning activities.

Feasibility: Currently measured in most countries, although measurement of completion could be improved.

2) Exposure to a breadth of learning opportunities across all seven domains.

Rationale: In Phase I, the task force identified seven domains of learning necessary for success in the 21st century. An even broader set of competencies is necessary at the national and local levels; however, the task force recommends that national curriculum, teacher training, and potentially instructional practices, be mapped, where possible, in relation to the seven domains.

Feasibility: No current measure exists at the global level to track the breadth of learning opportunities, and thus a new measure would need to be developed.

3) Early childhood experiences that promote development and learning in multiple domains.

Rationale: The early childhood years are critical to later learning and development. Entry to primary school is a key milestone in a child's learning trajectory, and measuring early development and learning across multiple domains at this point or in the years before entry to primary can help inform improvements in pre–primary education, health, family services, and other sectors serving young children. Since child development is influenced by multiple domains, a holistic measure across several domains is the best way to capture

learning at this stage. This typically includes aspects of learning related to five of the seven domains: physical well-being, social and emotional, literacy and communication, learning approaches and cognition, and numeracy and mathematics.

Feasibility: At present, several countries and regions use measures of children's learning at entry into primary or in the years immediately before. While there is no single measure that has been adopted globally, there are options for both country-level and global measurement, which are outlined in greater detail below. To reach global coverage with one measure, additional validation and scaling are needed.

4) The ability to read a variety of texts.

Rationale: Children and youth must be able to communicate in their mother tongue and in the primary language of instruction. The foundational skills necessary for learning to read are critical to functioning in modern society, in addition to the ability to comprehend and analyze complex texts through a variety of media. This area encompasses both primary and lower secondary levels.

Feasibility: Most countries have some measure of reading used at a national or subnational level, and some countries participate in internationally comparable assessments of reading skills.

5) The ability to use numbers and apply this knowledge to real-life situations.

Rationale: Children must be able to count and understand mathematical concepts, both to make informed choices in life and to pursue advanced learning in such disciplines as science, engineering, economics, research and technology. This measurement area encompasses both the primary and lower secondary levels.

Feasibility: Many countries have some measure of numeracy and mathematics used at a national or subnational level, and some countries participate in internationally comparable assessments.

6) An adaptable, flexible skill set to meet the demands of the 21st century.

Rationale: There are a variety of skills across the seven domains that children and youth need to succeed beyond reading and numeracy. A measure of these types of knowledge and skills administered in lower secondary school could include environmental awareness, collaborative problem solving, information and computer technology (ICT) digital skills, social responsibility or other subdomains.

Feasibility: Measurement of these skills is emerging, and some components are currently measured on a limited scale. However, a new measure or composite of measures would need to be developed to track progress at the global level.

To ensure that learning is measured and improved in an equitable way, information collected and reported within these areas should describe progress over time and across population groups in addition to average achievement levels in a country or region.

Equitable learning remains a key component of the LMTF's recommendations. Aggregate reports of learning outcomes at the country level can be useful for comparisons across countries, but more nuanced information is also needed to improve learning outcomes for all children. Measuring progress over time is necessary for recognizing those countries that have made significant progress in improving learning. In order to target improvements to the most marginalized children and youth, countries must also collect data

on socio-demographic factors that are associated with disparities in learning outcomes. At a minimum, countries should collect data on sex, urban/rural location, and family socioeconomic status.

Multiple methods should be considered when designing systems to assess learning opportunities and outcomes.

The task force agreed that rigorous assessment of learning can take multiple forms, including standardized assessments that are administered in one or more countries, internationally comparable assessments, national exams and assessments, and household-based surveys.

The areas of learning endorsed by the LMTF should be tracked by the global education community regardless of their incorporation into post-2015 global development goals.

While informing the post-2015 dialogue and agenda is one objective of the LMTF, the main focus of the task force is to collectively define global ambition on improving learning and propose practical actions to deliver and measure progress. These objectives can be achieved regardless of how education and learning are incorporated into the next round of development goals.

Countries must be supported in obtaining the financial, technical and political resources to measure learning and use the information to improve learning outcomes.

The task force recognized that a system of global measurement will only be effective in improving learning outcomes if there is a strong commitment to improve national and classroom-level assessment. The Measures and Methods Working Group proposed the establishment of a mechanism—such as a global, multi-stakeholder advisory group—to meet this objective. In Dubai, the task force agreed to explore the feasibility of using such a mechanism to support assessment in all forms, including in the six areas identified above and to help build national capacity for measuring learning.

Next, this report will detail the six areas of measurement—including the rationale and feasibility—and highlight some of the existing measurement efforts in each of the areas.

Six Areas of Measurement for Global Tracking

Assessments of learners' progress can be grouped into three broad categories: continuous classroom assessments, summative grades and examinations, and large-scale/system-level assessments. At the classroom level, assessments are ideally of a formative influence, where teachers experience firsthand their students' skill levels, and they generally serve to support and adjust teaching accordingly throughout the school year. However, in many classrooms around the world, assessment is primarily of a summative nature; that is, it entails providing grades at the end of the term. Examinations provide a signal as to how well learners have done at the main transition points within the school system and ensure that the intended curriculum is taught and learned; hence, they are geared toward making decisions about an individual student's progress through the education system. Finally, large-scale/system-level assessments are used for monitoring progress and providing policymakers with information on overall performance levels in the system, the changes in those levels, and related or contributing factors (Clarke 2012; UNESCO 2012).

National examinations are high-stakes assessments that have a direct effect on the students tested, as they are a tool to certify student achievement and determine who proceeds to the next level of education; they are therefore typically administered to all students at a specific grade/level. While they may have the unintended consequence of encouraging excessive attention to passing examinations rather than focusing on broader aspects of learning, many countries monitor learning outcomes through their national examination systems (UNESCO 2005). On the other hand, the assessments discussed in the following section are related to low-stakes, and to some extent no-stakes, assessment, with typically no consequences associated with student performance, and where many students perceive no personal benefit from the assessment testing experience (Wise and DeMars 2005).

The following section describes the rationale for recommending each area of measurement for global tracking, gives examples of existing data, and considers the feasibility of collecting these data in all or most countries. Indicators for each of the six areas will be decided upon by the task force and refined with the assistance of experts in these areas.

1. Access to and Completion of Learning Opportunities

Rationale: There are 57 million children of primary school age and 69 million of lower secondary school-age not currently enrolled in schools (UIS 2013). As defined by the UIS, out-of-school children at the primary level are those "in the official primary school age range who are not enrolled in either primary or secondary schools." These include:

- Children/adolescents who have dropped out; and those who are not, or have never been, enrolled in formal educational programs;

- Children/adolescents excluded from school due to gender, disability, conflict/disaster area, poverty, etc.;

- Children/adolescents who are yet to start school, but will eventually start school late; and

- Children/adolescents who are participating in non-formal or early childhood education programs.

The LMTF framework is structured to promote access plus learning. While the focus is primarily on learning, access is included in this framework so the millions of children still out of school are not excluded. For this reason, in addition to learning, global goals and measures should combine expectations regarding access and completion.

> *"Expanded access has little import unless it includes regular attendance, enables progression through grades at appropriate ages, and provides meaningful learning, achievement and completion." (Motala, Dieltiens and Sayed 2009)*

Monitoring access and completion in addition to learning outcomes is of crucial importance for being able to properly understand information on student achievement—as it has been observed, for example, that an expansion of access to schools could translate into lowering average levels of student achievement. Similarly, efforts need to ensure that average levels of achievement are not improved by inadvertently increasing exclusion. The potential misreading of evidence arises when school-based assessment data (i.e., assessments of learning conducted only in schools) are taken in isolation, without factoring the information on access levels into the analysis. Access plus learning means that increasing access (where needed) needs to be paired with actions pertaining to the improvement of learning. Some countries have experienced dramatic increases in access, but at the same time, learning levels have remained constant. For example, access to primary schooling in Mozambique, as tracked by the net enrollment rate, increased from 44% in 1990 to 90% in 2011, yet evidence from the Southern and Eastern Africa Consortium for Monitoring Educational

Quality (SACMEQ) has shown that learning levels have not changed much between SACMEQ II (2000) and SACMEQ III (2007).[1] It can, however, be said that if learning levels are at least constant, and access has increased, then the country is increasing the capabilities of its future population.[2]

Access and completion must be achieved with a focus on the most marginalized. Without disaggregating data for marginalized populations—including children with disabilities, those living in poverty or conflict situations, and in many cases girls and children living in rural areas—gains in overall access and completion can exclude the children who might benefit from education the most and could lead to further marginalization.

> *Consultees noted that access and completion for marginalized populations should not be ignored:*
>
> *"For children with disabilities, access still is an issue, and in most cases a significant issue. Any global indicators for enrollment, intake or completion are going to suffer from the 'low-hanging fruit' phenomenon, at least until the ratio gets very close indeed to 100%. Inclusion of children with disabilities, or indeed any given marginalized population, can't be ensured without some degree of intentional focus on that particular population."*

Existing data: Data on access (enrollment) and completion (educational attainment) are currently collected by the UIS at a global scale, even if, in some countries, there are certain limitations to the way age-specific data are recorded. While the UIS has conducted more comprehensive work on access, this report focuses on the components that are most representative and central to the LMTF's discussions.[3]

The net enrollment rate (NER) in primary (International Standard Classification of Education, ISCED, 1) and lower-secondary (ISCED 2) education is regularly produced at the UIS, to which 75% of countries report some data on enrollment, and 60% report enough data for producing the NER (UIS 2010). The inputs for the calculation of this indicator require countries to report, for each education level, their enrollment by single years of age and the population of the age group.[4] The data on enrollment generally come from school registers, school surveys or census reports on enrollment by age (UIS n.d.).

The net intake rate (NIR) for primary education is also regularly produced by the UIS; however, since 2000, only 66% of countries have provided sufficient data to report on the rate at least once, and 40% of countries have provided sufficient data to report on the NIR up to four times. The production of this indicator requires countries to provide information on the new entrants to the first grade of primary by single years of age, in addition to the population of primary school-entrance age. The data for this generally come from school registers, school surveys or census reports on new entrants by age (UIS n.d.).

There are many indicators of completion that are being readily used. It is, however, crucial to differentiate between ratios aimed at measuring volume from rates that measure the proportion of a given population that has actually completed or is likely to complete a given educational level.

Further analysis would be needed to adjust for the amount of time it takes for a student to complete a given level by identifying those who complete that level "on-time," as well as students who complete that level regardless, regardless of timing. On the other hand, the EFA Global Monitoring Report has introduced the "expected cohort completion rate," a measure that es-

timates the likelihood of completion for a given cohort. For example, for primary, the expected cohort completion rate is the proportion of children of primary school starting age who are expected to complete primary school, including those who start late and repeat primary school grades.

Feasibility: All the indicators mentioned above are suitable for global tracking, and data are currently available. However, these indicators should be handled with caution. There has been extensive research on the advantages and disadvantages of the current indicators used for monitoring access and completion. The following discussion highlights some of the issues that need to be taken into account.

In addition to measurement errors, there may be other uncertainties in the data, particularly when data on enrollment by age are required or when data are based on population projections. As explained by Lewin (2011), without knowing how children are distributed across grades (e.g., within the primary level), the NERs may not be an appropriate indicator of "timely" coverage, and changes in values can have more than one cause. Data on timely entry (net intake) in grade 1 may also be unreliable because of poor data on enrollment by age.

The NER compares the total enrollment of children of the appropriate age for the cycle with the level's school age population. While the rate excludes those over the nominal age for the last year of a given level, it does not factor in those who are over-age in the lower grades (Lewin 2011). The interpretation of metrics that are age-specific and time bound may create confusion. For example, there is no scientific evidence on whether a country with a total completion rate of 80%, and a timely completion rate of only 60%, is better or worse off than one with a total completion rate of only 75% but a timely completion rate of 65%.

2. Exposure to a Breadth of Learning Opportunities across All Seven Domains

Rationale: In Phase I of the LMTF, the task force identified seven domains of learning necessary for success in the 21st century: physical well-being, social and emotional, culture and the arts, literacy and communication, learning approaches and cognition, numeracy and mathematics, and science and technology. While tracking learning outcomes in all seven domains at the various stages of learning (early childhood, primary, and lower secondary) is not pragmatic at a global or even national level, it is possible to measure children's exposure to learning opportunities in these areas. This could help ensure that the focus of learning remains on a broad spectrum of domains, even if learning outcomes in only a subset of the domains are measured at the national and global levels. The subdomains within the seven domains framework could be adapted and expanded per the national context to account for differences in national curricula and priorities.

> *One working group member highlights the need to adapt the seven domains to the national context:*

> *"I agree that the development of a new instrument to record the inclusion of the seven domains in national curricula policies would be valuable. This would give us a good picture of the situation and an idea of how feasible a global set of measured domains really is. I feel that the subdomains would necessarily have to be included in this exercise as they would give the extra granularity required for some of the developing countries."*

The task force recommended further investigation into a tool or survey that would capture this information at the global level. The task force also recommends that national curriculum, teacher training, information on textbooks and materials, and, potentially, instructional practices be mapped in relation to the seven domains where possible. This measure, when combined with learning outcome measures, could provide valuable information about the inputs necessary to achieve a more holistic view of learning outcomes.

Examples of existing measurement efforts: There are related initiatives under way that will be examined by the Implementation Working Group in forming their recommendations, including:

International Review of Curriculum and Assessment Frameworks Internet Archive (INCA):
http://www.inca.org.uk/country_archives.html

Provides descriptions of government policy on education in high-income countries (Australia, Canada, England, France, Germany, Hungary, Ireland, Italy, Japan, South Korea, the Netherlands, New Zealand, Northern Ireland, Scotland, Singapore, South Africa, Spain, Sweden, Switzerland, the United States and Wales) and makes particular reference to the curriculum, assessment and initial teacher training frameworks in place.

Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS) Encyclopedias:
http://timssandpirls.bc.edu/timss2011/encyclopedia-timss.html

TIMSS Encyclopedia provides an overview of the national contexts for mathematics and science education in the countries participat-

ing in TIMSS 2011. Nearly all the TIMSS 2011 countries and benchmarking participants prepared a chapter summarizing the structure of their education systems, mathematics and science curricula and instruction in primary and secondary grades, teacher education requirements, and assessment practices.

PIRLS Encyclopedia provides an overview of the national contexts for teaching and learning reading in the countries participating in PIRLS 2011. Nearly all the PIRLS 2011 countries and benchmarking participants prepared a chapter summarizing the structure of their education systems, reading curricula and instruction in primary grades, teacher education requirements, and assessment practices.

World Data on Education:

http://www.ibe.unesco.org/en/services/online-materials/world-data-on-education.html

Produced by the UNESCO International Bureau on Education, this database contains detailed and systematized information on education systems of 161 countries worldwide, with a particular emphasis on curricula and curriculum development processes.

Feasibility: No such instrument currently exists at the global level to track the breadth of a student's learning experiences across domains, and therefore a new instrument would need to be developed. As a starting point, information could be gathered using simple checklist (Yes/No) or an ordinal scale (e.g., with 4 categories) requesting countries to confirm the presence (or extent) of the domains and subdomains of learning by education level in:

(i)   teacher initial training programs;
(ii) teacher in-service training programs;
(iii) national curriculum and assessment framework; and
(iv) content of educational materials (e.g., textbooks).

Simply examining the national policies does not provide information on how the policies are implemented. However, gathering information on these policies could be an important first step in developing and scaling up more robust tools to measure the quality of learning opportunities to which children are exposed.

3. Early Childhood Experiences that Promote Development and Learning in Multiple Domains

Rationale: Entry to primary school is a key milestone in a child's learning trajectory. There are multiple factors that result in children succeeding upon entering school: Children must enter school with the cognitive and noncognitive skills and experiences needed to be successful; schools and teachers must be equipped to provide welcoming and stimulating learning environments; and families must be ready to help their children make the transition to school and support their learning. While the focus of this area of measurement is on the child's development at or around school entry, all three factors should be in place to support children entering school so they will be ready to succeed.[5]

Children's development before school entry has a profound and lasting effect on school achievement; rather than "catching up" once school starts, children who do not have the necessary skills and knowledge before starting school tend to fall farther and farther behind as the school years progress. The role of early childhood development in school achievement and comple-

tion has been established through scientific evidence that cuts across disciplines. Together, this research shows that learning trajectories that persist throughout schooling begin in the early years, and are shaped by physical, social/emotional and cognitive development.

These domains work together to promote learning, and each domain plays a critical role in influencing children's learning. The LMTF recommends that measurement of children's development at school entry should include at least five domains of development: physical well-being, social and emotional, literacy and communication, learning approaches and cognition, and numeracy and mathematics. Like learning at other stages of life, children's learning at the time of school entry reflects both those skills and types of knowledge that are common across countries, and those that are more sensitive to the specific goals and approaches within a given country. Common competencies include socially-appropriate behavior with peers and adults and emergent communicative, literacy and mathematics skills. How these specific competencies are demonstrated may vary based on countries' goals and priorities for early childhood. Data on learning at school entry can be used for several purposes: to identify areas for interventions in the years before school; to shape and inform curricula and teacher training efforts; and to ensure that there is equity in access and quality of preschool education and health/nutrition services. Reporting data by domains may be especially useful for identifying areas where additional services are needed in the early years to promote learning at school entry.

While there are multiple measures of child development and learning available for this age group, and some may be useful for global comparisons, the LMTF recommends that countries approach learning at school entry using measures aligned closely with national priorities. In some countries, measurements are administered within the first few months of primary school. In others, in addition to covering multiple domains of children's learning, the measurement of learning at school entry should also reflect countries' clear delineation of goals and priorities for early childhood learning. For this reason, those consulted in the LMTF process suggested that learning at school entry be captured through nationally defined assessments rather than relying on one assessment used by every country. As outlined in greater detail below, there are several measures available now that offer valid and reliable assessments of children's competencies at school entry or in the years before that are used in multiple countries. These measures can be adapted and adopted by countries to align with national priorities.

It is important to note that the purpose of assessment at this level is to help teachers, schools and education systems understand the abilities and needs of students as soon as possible so that they can meet the child's educational needs and identify any problems in development as early as possible. The purpose should not be to delay primary school entry for individual children or to otherwise use assessment in a high-stakes manner.

Examples of existing measurement efforts: There are several tools available that measure children's learning at the start of school and their "readiness" to learn within instructional settings. What follows is not an exhaustive list but offers some ideas on how existing efforts have approached large-scale measurement of learning at school entry. These tools are generally administered by teachers and administrators in classroom settings; sometimes a portion or the entirety of the instrument is completed by children's parents.

Tools used by countries or regions to measure learning at school entry include:

- The Early Development Instrument (EDI), developed by the Offord Centre for Child Studies at McMaster University (Janus and Offord 2007), measures readiness to learn in school at the population level. Teachers administer the EDI in preprimary and primary programs after a child has attended school for approximately 4 months. To date, it has been used in 24 countries. Teachers rate the individual children on each of the questions included in the instrument and the data are reported and analyzed in aggregate (therefore, EDI is not used as a diagnostic for individual learners). Of the topics included in the EDI questionnaire, the following are the most closely related to the seven domains of the LMTF: physical health and well-being; social competence; emotional maturity; language and cognitive development; and communication skills and general knowledge. The EDI is a proven valid and reliable measure that contains concurrent and predictive associations for school achievement in mathematics and reading.

- The East Asia- Pacific Early Child Development Scales (EAP-ECDS) were recently developed by the Early Childhood Development, Education and Policy Group within the Faculty of Education at the University of Hong Kong. The project to develop and validate the Scales has been supported by UNICEF and the Open Society Foundations and is being overseen by the Asia-Pacific Regional Network for Early Childhood. The Early Learning and Development Standards from 7 countries in the East Asia and Pacific Region were used to develop the items on the scale, which are considered to be appropriate for children aged 3-5. Based on the results of pilot studies conducted in China, Fiji, and Mongolia in 2011, the latest version of the EAP-ECDS includes 7 domains (Approaches to Learning; Cognitive Development; Cultural Knowledge and Participation; Language and Emergent Literacy; Motor Development; Health, Hygiene, and Safety; Socio-Emotional Development) and 85 items. The scales are currently being validated on representative samples in Cambodia, China, Mongolia, Papua New Guinea, Timor-Leste and Vanuatu.

Examples of efforts at the national level include:

- The evaluation efforts for Korean NURI-Curriculum, based on the Early Childhood Assessment Scale, assesses five-year-old children's knowledge, skills and attitudes in areas determined by the national curriculum, which include physical activities and health, communication, social relationships, art experiences, and nature. NURI for five-year olds has been implemented in approximately 20,000 preprimary programs across the country.

- The Hong Kong Early Child Development Scale assesses children in preprimary programs of age three to six years in the following domains: physical fitness and health, language, self and society, early mathematics, science and technology, and arts.

- The School Readiness Instrument (SRI), administered in India, is a tool used for program evaluation, and providing indicators of foundational skills; therefore, it covers selective domains within cognitive and language development. Unlike the other assessment tools listed here that teachers, administrators, or parents administer to children, a researcher who may not necessarily be familiar with the child administers the School Readiness Instrument. Therefore, the researcher may not be able to accurately assess the child's social and emotional development if the child does not have an opportunity to demonstrate the competencies during the observation period. Nevertheless, SRI is being complemented by the Assessment of ECE Center/Anganwadi Center, which provides information on the quality of programs by making regular observations in classrooms. This includes a measure of the frequency of opportunities observed for helping children develop personal and social skills and competencies.

- Work Sampling for Head Start (published by Pearson) is used by some federally funded preprimary programs in the United States to assess children's abilities using samples of their classroom work and an observational checklist. The instrument measures learning and development in areas that correspond to all seven LMTF domains. This

information is used both by classroom teachers and can be aggregated to be useful at the policy level (Shamblott 2012).

- The IDB has also been developing Regional Indicators of Early Childhood Development for Latin America (PRIDI). Currently, four countries (Costa Rica, Nicaragua, Paraguay, and Peru) have tested instruments and data have been collected on Early Childhood Development Indicators for samples of 2,000 children per country.

Feasibility: The level at which learning and development in early childhood is assessed may depend on the country's availability of preprimary school programs and on the age at which compulsory schooling begins. In countries where the availability of and participation in preprimary programs (ISCED 0) is low, it may be more appropriate to measure learning and development in the first grade of primary (ISCED 1), or whenever compulsory schooling begins. Alternatively, a school readiness component added to household surveys, such as the Citizen-Led Assessments, or measures of child development, such as those included in the Multiple Indicator Cluster Survey (MICS) conducted in households, could be used for all children, regardless of whether they are enrolled in preprimary programs.

4. The Ability to Read

Rationale: Children and youth must be able to communicate in their mother tongue and in the primary language(s) of their school and society. Communication in one's mother tongue requires individuals to have knowledge of vocabulary, functional grammar and the functions of language and script (European Commission 2006b). Reading is a complex process, in which individuals activate a wide range of competencies when thinking about, monitoring and adjusting their reading activity for a particular task or goal.

Reading should not be seen as restricted to the ability to decode text, but is related to individual's knowledge of words, grammar, linguistic properties and properties of the script, as well as other metacognitive competencies, such as the awareness of and ability to use a variety of appropriate strategies when processing text (OECD 2009). As children develop their reading skills, they should also develop their ability to read critically to gain a deeper understanding, by locating, evaluating and integrating information. Reading skills have become necessary for continuous learning in a world that is rapidly changing. The importance of reading comprehension is highlighted in today's information age, in which the ability to read easily has become a critical skill. Reading, which entails understanding, is a key outcome of primary education, and one of the core elements in the definition of the level according to ISCED 2011.

Developing the ability to comprehend and analyze texts delivered through different types of media are critical to functioning in modern society and require a set of foundational pre-reading skills. Foundational skills—such as being able to recognize letters, decode words and read short sentences—have mostly been emphasized in the early grades, while reading comprehension has been the focus in the later grades of primary (ISCED 1) and beyond primary education.

> *There was broad support from the working group and consultations that reading skills should be measured at several points during the education cycle, focusing on different aspects of reading.*

Measurement of foundational skills can enable countries to determine if their children are learning to read independently. At the same time, measuring reading comprehension in the later grades would assess

children's ability to understand and gain subject matter knowledge, enabling them to practice "reading to learn." Measurement at different grades or ages would give countries the opportunities to take necessary action to target learning interventions and improve students' performance. Countries may choose to assess reading skills for many reasons: policy dialogue, informing instruction, monitoring progress, assessing the success of a reading program, and as a baseline to inform program design.

> *There was broad support from the consultation that measurement of foundational reading skills should be done at least by grade 3, mostly as a way to make it possible to improve performance by the end of the primary education cycle.*

"Learning to Read" Assessments—Measuring Foundational Reading Skills

Examples of existing measurement efforts: Assessing foundational skills typically takes place in the early grades of primary education. Since 2005, at least 61 developing countries have measured foundational reading skills of children in the early grades of primary schooling at least once.[6] These assessments are typically composed of several exercises, which are administered in increasing difficulty from simple oral language development skills such as sound awareness to more complex tasks related to reading comprehension. These assessments generally include tools for assessing phonological awareness, sound-print correspondence,[7] receptive and expressive vocabulary, decoding, reading fluency, and oral or listening comprehension.

Even though many of the assessment instruments developed for measuring foundational reading skills have common features, they are not intended to produce in-

ternationally comparable results on an equivalent metric—not only because of the different languages and scripts, but also because of the specific characteristics of the settings and data collection methods.

Assessments of foundational reading skills used in multiple countries include:

- The Early Grade Reading Assessment (EGRA) is an individually administered oral student assessment that has been conducted in more than 40 countries and in a variety of languages. Administered in grade 2 or 3, the interview lasts approximately 15 minutes and includes timed assessments of letter naming, letter sound knowledge, phonemic awareness, pseudo-words and familiar words, oral reading fluency and untimed segments, including reading comprehension, listening comprehension and dictation. EGRA is a school-based assessment.

- The Annual Status of Education Report (ASER) is a citizen-led assessment that has been conducted every year since 2005 in all rural districts of India (over 575 districts), in 20 languages. ASER has also been administered yearly since 2008 in urban and rural areas of Pakistan in four languages. In 2012, ASER covered 136 rural districts and 6 urban districts in Pakistan. It is administered in households to children ages 5 to 16 years old. The reading components include letter naming, reading simple words, sentences and a short paragraph—in some years in the Indian ASER, a reading comprehension module was added that asked the child to read a text, retrieve factual information, and provide the answer orally or to read a paragraph, its questions, and provide the answer to the questions orally. Additionally, English reading skills were assessed for all selected children in 2007, 2009 and 2011, regardless of their mother tongue. ASER in India reaches between 600,000 and 700,000 children each year. The ASER survey reaches a representative sample of children from every rural district in India. It includes children who are enrolled in government schools, in private schools and in other kinds of schools and also children who are currently not in school.

- Uwezo is also a citizen-led assessment of a nationally representative household survey that has been conducted annually in Kenya, Tanzania and Uganda since 2009. Like Uwezo, which was inspired by ASER in India, other efforts are under way in Mali (called Beekungo) since 2011, and more recently in Senegal (called Jangandoo) for children age 6 to 16 years. The foundational reading skills assessed include letter naming, reading simple words, sentences and a short paragraph, which is paired with simple comprehension questions.[8]

- The Literacy Boost Assessment, developed by Save the Children, is administered in schools to children who are 6 years or older in grades 1 through 4 across 25 project sites in 19 developing countries across Africa, Asia, Latin America and the Caribbean.[9] The foundational reading skills assessed include concepts of print, knowledge of the alphabet, single word reading, timed oral reading fluency and untimed accuracy, and reading comprehension questions.

- SICOLE-R-PRIMARIA is a school-based assessment administered to children age 6 to 11 years in Chile, Guatemala, Mexico and Spain. An adapted version of the test has also been administered in a smaller sample in Ecuador. The foundational reading skills assessed include speech perception, letter naming, pseudo-words and word naming, phonemic awareness, homophone comprehension tasks, morphological comprehension tasks, and other reading comprehension exercises.

Feasibility: Reading acquisition is highly linked to the specific characteristics of language and script, and hence, assessment instruments are adapted or even redeveloped to be relevant for each combination of language, script and content. Therefore, the characteristics and varying complexity of each language's orthography presents a barrier to a truly global metric that captures the full range of skills relevant at the foundational stage.

However, it is possible to identify students who did not complete any exercise presented to them—or in other words, those who have obtained zero scores.[10] For example, the number of children who cannot recognize a single letter or decode a single word may be compared.[11] Therefore, countries can report on the percentage of children who are not able to decode a single word by a certain school grade or age group. This statistic may be potentially comparable across different studies and countries, provided that the assessment tool is administered to a representative sample of the children tested at a given grade. It would also be possible for each country to set its own target; for example, to decrease the number of children who cannot read in the early grades (or those of the appropriate age) by a certain number (or percentage) by a given date.

Countries could also use instruments for assessing foundational reading skills to report on the percentage of children who are making sufficient progress in the early grades according to nationally-defined criteria. This means that countries that do not have clearly defined reading benchmarks in early grades would have to do so. Therefore, the percentage of children who are improving would not be comparable internationally, but would be based on each country's own definition of "sufficient progress." It would also be possible for each country to set its own target for improvement; for example, to increase the number of children who are making sufficient progress in the early grades by a certain number (or percentage) by a given date.

Measuring foundational reading skills for out of school children is captured in some countries through household surveys (as in the cases of ASER and Uwezo noted above), and stakeholders within countries must judge the cost-effectiveness of collecting such information. While ASER is not a longitudinal study, it is ad-

ministered yearly. Efforts like ASER would be one way of tracking the progress of children of a given age or grade level from year to year, regardless whether they are in school or not.[12]

"Reading to Learn" Assessments—Measurement of Reading Comprehension Skills

Much effort has been devoted to measuring reading in the later grades of primary (ISCED 1) and in lower secondary education (ISCED 2). Based on UIS estimates, more than 150 countries conduct some type of national assessment or examination and/or participate in international studies of student achievement.

Whether designed to be administered in schools or households, reading comprehension tasks address broad cognitive strategies such as the ability to retrieve or locate information within a text, form a broad understanding of or interpret a given text, or reflect on and evaluate the content and form of a text (IEA 2011; UIS 2009; OECD 2009). Each assessment typically develops a framework that identifies and defines the cognitive strategies or comprehension processes that are necessary for achieving full understanding of texts. Although the tasks may vary by assessment, they generally include several types of comprehension processes; and the test questions, each addressing one of the processes, enable respondents to demonstrate a range of abilities and skills in constructing meaning from written text (IEA 2011).

Examples of existing measurement efforts: Reading is measured through a number of different initiatives at the global level. Approximately 80 countries are currently participating in regional or international assessments, where scales have been established and the different levels have been attached to definitions that have been statistically validated. While many countries have developed their own national assess-

ments of reading, the following assessments have been used cross-nationally:

- International Association for the Evaluation of Educational Achievement's (IEA) Progress in International Reading Literacy Study (PIRLS), and a complementary version called pre-PIRLS, is a school-based assessment that is conducted every five years. In 2011, 49 countries participated in PIRLS, which assessed students in the grade that represents four years of schooling counting from the first year of ISCED 1; optionally, countries could choose to assess in later grades (up to grade 6), or to use the less difficult pre-PIRLS version.

- The Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), coordinated by UNESCO's regional Bureau for Education in Latin America and the Caribbean, is currently in its third study of student achievement among 15 countries in Latin America. It assess students in grades 3 and 6.

- The Programme d'analyse des systèmes éducatifs de la Conférence des ministres de l'Education des pays ayant le français en partage (PASEC/CONFEMEN) is a school-based assessment that has started its next cycle (2013–16). It assesses student performance in grades 2 and 6 in 10 countries in Central and Western Africa. In its previous cycle, between one and three countries in francophone Africa were assessed each year from 1991 to 2010.

- The Southern and Eastern Africa Consortium on Measuring Educational Quality (SACMEQ) is a school-based assessment administered in 14 countries in Southern and Eastern Africa to students in grade 6; SACMEQ has completed three rounds: 1995, 2000 and 2007.

- The Program for International Student Assessment (PISA) assesses students' abilities to understand, use, reflect on and interpret texts (print and electronic) among 15-year-olds who are enrolled in schools in grade 7 or above. PISA is implemented every three years, and in addition to reading, assesses mathematics and science with a focus on students' content knowledge, skills and competen-

cies reported on proficiency levels. Over 70 countries participated in PISA 2009 and results from the current cycle of PISA 2012 will be published in December 2013. A pilot trial called PISA for Development is currently being implemented with a select group of countries to explore ways to increase the relevance of PISA for contexts in developing countries while maintaining international comparability. During the pilot stage, PISA for development will also consider an approach and methodology for including out-of-school children in the assessment.

For both national and international assessments, the UIS Observatory of Learning Outcomes aims to catalogue key information on the design and implementation of learning assessments to yield standardized, comparable information on the availability of assessments and their main characteristics. While data on student assessment are increasingly available—whether administered in small samples, at the national or regional level—cataloguing country portfolios or summaries of assessment subjects, their results, and providing detail on assessment design will serve to provide a resource for those who wish to better understand the range of assessment approaches and facilitate the broader use of the results.

Feasibility: Although all the above-mentioned assessments are administered in school (potentially with the exception of PISA for Development), they differ from one another on several key features, including purpose and proximity to the curriculum and content (Stephens and Coleman 2007). PISA is not directly based on curriculum, but on students' knowledge, competencies and skills, so the assessment focuses on students' ability to reflect and apply their knowledge and experience to real-life situations, rather than only on content learned in school (OECD 2009). The regional assessments—LLECE, PASEC and SACMEQ, as well as IEA's studies including PIRLS—are based on the curriculums and the wide range of interests in participating

countries. Focusing measurement efforts in schools may be pragmatic, not only because it is economical, but also because the primary aim of large-scale assessments is to influence education policy and school systems; however, this continued focus on measuring learning within schools may lead to further exclusion of out-of-school children and youth. Household-based surveys can provide information on all children and youth, including those who are not enrolled in school, are enrolled in nonformal schools and private schools, or are absent on testing days. However, the measures used in household surveys must be valid and reliable in order to be useful—in the past, literacy has been measured via household surveys in many countries through simply asking members of the household whether they are literate (Schaffner 2006).

It is assumed that all education systems aspire to the goal of having all their children reading competently, that is, understanding and not only decoding texts—moving from "learning to read," to "reading to learn" as they progress through the school cycle. For this reason, reading comprehension assessments generally take place in school, although some household-based assessments of reading comprehension are administered to youth and adults.[13] If the goal to have all children reading competently by a certain level or age is purposefully defined broadly, it allows each country to pick and set benchmarks relative to its national curriculum and the assessment tools that it currently uses for measuring reading.

Given the wide availability of assessment tools, countries may be able to report on the percentage of children at the end of primary schooling achieving a certain level of reading proficiency. Country-developed metrics have the potential for ensuring relevance if they are closely linked to the national policy priorities and agenda. At the same time, it should be clear that

information generated using different metrics (even when inspired by a common set of guidelines) cannot be aggregated at the global level. Comparing results, scores, or distribution among levels from different metrics can be misleading.

If there is interest in having a common global metric on reading comprehension, several approaches could be explored, each with advantages and limitations:

1. Expanding the scope of existing internationally comparable studies, which would rely on methods that have already been tested and validated (and thus save developmental costs). Limitations may stem from the group of countries where the metric has been validated; to the extent that more low-income countries participate in existing comparable measurement efforts, the risks of having significant floor effects would be minimized. Floor effects occur when the test is too difficult and many scores are near or at the minimum possible score (Ary et al. 2006). While studies with significant floor effects still generate useful information—for example, they may indicate that a high percentage of the targeted population has not achieved even basic skill levels—they would also limit the further analytical possibilities of the data.[14] Both limitations become less of an issue when participating countries collaborate in terms of determining the overall design of the study. Additional limitations may also stem from proprietary rights to assessment instruments and the idea of international statistics as public goods.[15]

2. Linking existing studies that have been implemented and refined is another way to generate comparable information. While existing regional initiatives address specific policy agendas and encourage collaboration, linking may be difficult, especially given the diverse institutional strengths of the different players, the high degrees of developmental work required, and the additional in-country costs.

3. Developing a new metric has a clear advantage in that it can build on the knowledge base derived from existing regional and international assessments and build upon this knowledge base. The main limitation to developing a new metric is the significant amount of developmental work required. Achieving global comparability, in areas such as reading and mathematics, demands a significant and sustained effort. This is an ongoing developmental and dynamic process especially as new technologies and techniques could improve the quality of the assessment.

Assessments that target specific age cohorts, rather than certain grades, may test students who are enrolled at different stages of the education system—different grades or education levels. In some countries, students have a high probability of being enrolled in school past primary education, and at the grade expected given their age—but this is not the case where there is no universal access, and late entry or grade repetition are common phenomena. Additionally, in many countries, children enrolled in schools are typically more advantaged (socially and/or economically) than those who are not—therefore, if only based on children who are in school, the assessment would not yield results that are generalizable to the targeted age cohort. Therefore, assessments are required for children who are not participating in the formal schooling system in many countries to provide relevant and usable policy insights about the whole cohort. Beatty and Pritchett (2012) suggest that for learning goals to capture the progress of all children, they should be formulated for an age-specific cohort, and not just for children in school; this would allow policymakers to focus on improving competencies and capabilities of the entire population. For these reasons, the LMTF recommends, at the very least, that reading data be paired with access data to prevent misrepresentations of the characteristics of the population.

5. The Ability to Use Numbers and Apply this Knowledge to Real-Life Situations

Rationale: Children must be able to count, understand and use mathematical concepts both to make informed choices about different areas of life and to pursue advanced learning in such disciplines as science, engineering, economics, research and technology. Mathematical reasoning possesses value of its own and helps to train the mind for critical and disciplined inquiry. There has been increased awareness that a person's effectiveness as a citizen and success in the workplace are greatly enhanced by knowing and being able to use mathematics (IEA 2011). Numeracy is another central expected outcome of primary education. Thus, it is also one of the core elements in the definition of the level according to ISCED 2011. Similarly to reading skills, numeracy and mathematics can be thought of in two dimensions—the first, is foundational, focusing on number concepts, operations, geometry and patterns (as defined in LMTF 2013); and the second, where a student's thinking process is assessed when engaging in mathematical content. Early mathematical knowledge is a primary predictor of later academic achievement, and future success in mathematics is related to an early and strong conceptual foundation (Duncan et al. 2007). It is concerning to note that out of the eight developing countries that participated in TIMSS between 1995 or 1999 and 2007, performance levels in mathematics have fallen for six of them (Beatty and Pritchett 2012).

Assessments of foundational numeracy skills, similarly to those of reading, are administered both in and out of school, and are typically composed of several exercises that are administered in increasing difficulty from simple number identification to shape naming and simple arithmetic operations such as addition or division. It is important to note and to control for the potential interference of reading and writing skills in

test results—that is why some tests are designed using a fully oral approach. These assessments may be used to establish a baseline national snapshot, monitor progress, inform instruction, or inform policy. In the later years of primary schooling, assessment of numeracy skills focuses on the application of mathematical knowledge—such as operations, solving problems, interpreting data, understanding data displays and reasoning. While the cognitive dimension of numeracy assessments is measured in primary and lower secondary schooling, the focus in the later years shifts from an emphasis on factual knowledge, concepts and procedures to a more targeted focus on the ability to reason in unfamiliar situations, complex content and multistep problems (IEA 2011).

Work on the measurement of numeracy at a global level is somewhat less advanced than that on reading skills. However, some organizations have already set goals to improve numeracy skills in the early grades. For example, GPE has set the objective to dramatically increase the number of children learning and demonstrating mastery of basic numeracy skills by grade 3 (GPE 2012).

There was broad support from the working group and consultations that numeracy skills should be measured at several points of the education cycle. While the task force recognizes that "a strong foundation in mathematics during the early grades is crucial for success in mathematics in the later years" (RTI International 2009), the focus will be on tracking children in the later years of primary (or in lower secondary) school in their knowledge of mathematics, their ability to apply that knowledge and reason in unfamiliar situations, complex content and multistep problems (IEA 2011).

Examples of existing measurement efforts: There are many efforts aimed at assessing children's abilities

in numeracy and mathematics, in the early grades, in primary and in the later grades. Cross-nationally administered assessments of foundational reading skills also include a numeracy component, namely:

- Early Grade Math Assessment (EGMA) is an individually administered oral assessment of foundational mathematics skills. The assessment includes counting, number identification, quantity discrimination measures, missing number measures, word problems, addition and subtraction, shape recognition and pattern extension. EGMA was piloted in Kenya in 2009, and is currently being (or has been) implemented in 12 countries to students in grades 1 to 3.[16]

- The ASER is administered yearly in households in India and Pakistan, where numeracy tasks include number recognition, or the ability to recognize two-digit numbers, the ability to do basic arithmetic such as subtraction (with borrowing) and division of three-digit numbers by a single number.

- Uwezo, conducted annually in Kenya, Tanzania and Uganda, includes counting and matching, number recognition, quantity discrimination of double digit numbers less than 100, simple additions (without carrying), subtraction of two-digit numbers without borrowing, multiplication by a number not more than 5 and the total must be less than 25, division by a number less than 5 and the number being divided should be less than 5.

- Numeracy Boost, developed by Save the Children, has been launched in Bangladesh and Malawi in 2012 and more recently in Tanzania. It includes counting, one-to-one correspondence, number identification, skip counting / counting backwards, number discrimination, missing numbers, timed and untied versions of addition and subtraction, word problems, geometry, and measurement and time.

There have also been cross-nationally administered assessments of numeracy skills in the later grades of primary and lower-secondary:

- International Association for the Evaluation of Educational Achievement's (IEA) Trends in International Mathematics and Science Study (TIMSS), and a complementary version called TIMSS-Numeracy, is a school-based assessment of mathematics and science, administered in grades 4 and 8, and is conducted every four years.[17] A total of 52 countries thus far have participated in TIMSS, which also includes additional questionnaires for students, parents/caregivers, teachers, and school principals to generate information on home- and school-level experiences affecting the process of learning to read. TIMSS assesses grade 4 and 8 students in both subject matter knowledge (numbers, algebra, geometry and data and chance at each level), and their mathematical thinking process, which include knowing, applying and reasoning mathematical content.

- The above-mentioned regional assessments— SACMEQ (grade 6), PASEC (grades 2 and 6), and LLECE (grades 3 and 6)—also measure numeracy and mathematics in primary school.

- PISA measures students' capacity to formulate, employ and interpret mathematics in a variety of contexts. Students who perform at the baseline level of PISA proficiency in mathematics can employ basic algorithms, formulae, procedures or conventions and they can interpret and recognize situations in unfamiliar contexts that require no more than direct inference (OECD 2009). PISA for Development will explore ways that component skills in mathematics can be assessed to better describe abilities of students that may perform below the baseline proficiency.

Feasibility: As in the case for reading, it is assumed that all education systems aspire to the goal of enabling children to not only know mathematical concepts but also be able to use them in developing mathematical skills and applying them to daily life. Generally drawn on curriculum standards, mathematical skills are assessed mostly in schools; nevertheless, there are also numeracy components in tests of

the adult population—like the UNESCO Institute for Statistic's LAMP (15 years and older population) and OECD´s PIAAC (16–65 population)—that are administered in households.

Given the availability of assessment tools, countries may be able to report on the percentage of children at the end of the education level (primary and lower secondary) achieving a certain level of mathematical proficiency. Only when relying on a common metric can data be compared or aggregated to provide a global picture. The potential development of a common global measure on numeracy could explore several approaches:

1. Expanding the scope of existing internationally comparable studies. For the advantages and limitations of this approach, see the previous section on reading comprehension.

2. Linking existing studies for numeracy, such as TIMSS, LLECE, PASEC, and SACMEQ, may be possible but has not been explored by agencies involved in these studies to the same extent as reading.

3. Developing a new instrument that defines mathematical competence, and takes into account the considerations outlined in subsequent sections of the report. For the advantages and limitations of this approach, see the previous section on reading comprehension.

For more details on the feasibility of tracking numeracy skills, refer to the previous section on reading.

6. An Adaptable, Flexible Skill Set to Meet the Demands of the 21st Century

Rationale: There are a variety of skills across the seven domains that children and youth need to succeed beyond reading and numeracy. A measure of these types of knowledge and skills administered in

lower secondary school could include environmental awareness, collaborative problem solving, ICT digital skills, social responsibility and other subdomains.

The UN Secretary-General's Global Education First Initiative describes global citizenship in the following way: "The world faces global challenges, which require global solutions. These interconnected global challenges call for far-reaching changes in how we think and act for the dignity of fellow human beings. It is not enough for education to produce individuals who can read, write and count. Education must be transformative and bring shared values to life. It must cultivate an active care for the world and for those with whom we share it. Education must also be relevant in answering the big questions of the day. Technological solutions, political regulation or financial instruments alone cannot achieve sustainable development. It requires transforming the way people think and act. Education must fully assume its central role in helping people to forge more just, peaceful, tolerant and inclusive societies. It must give people the understanding, skills and values they need to cooperate in resolving the interconnected challenges of the 21st century."

Examples of existing measurement efforts: Some aspects of these skills are currently measured, but not at a global scale, including:

• *Citizenship:* IEA has developed a series of international studies on citizenship since 1971. The latest version was the 1999 International Civic and Citizenship Education Study (ICCS). This study, conducted in grade 8, and including regional modules, can become the foundation for a more global effort in years to come. A new round of this study is currently under preparation and will be undertaken in 2016.

• *Learning approaches and cognition:* TIMSS and PISA collect information on a student's background and the school environment; additionally, TIMSS collects information on teachers. TIMSS's Teacher

Questionnaires cover topics of academic preparation and certification, teacher assignment and induction, professional development, teacher characteristics, curriculum topics taught, class size, instructional time and activities, assessment and homework, computers, and Internet and calculator use (Hutchison and Schagen 2006).

- *21st-century skills:* The Assessment & Teaching of 21st-Century Skills (ATC21S) project examined multiple frameworks for 21st-century skills and developed a framework for assessment in two skill areas: collaborative problem solving and ICT skills (ATC21S 2012).

- *Social responsibility:* The Developmental Asset Lists contain items related to identity, social justice, caring for others, responsibility, social competence, etc. This measure has been translated in to 15 languages and has competency lists spanning from early childhood through post-primary (Search Institute 2006).

Feasibility: The idea for this measure is that it cuts across different domains (most likely social and emotional, learning approaches and cognition, science and technology) and could potentially be reported as a composite or index. The task force encouraged the Implementation Working Group to develop options that went beyond paper-and-pencil assessments to include computer-based and perhaps mobile-device-based assessments. Terms such as global citizenship, 21st-century skills, transferable skills and noncognitive skills have all been used to describe elements of learning in this area. There is currently no consensus in the policy or academic communities (and indeed, there is much debate) on what these terms mean and how they can be operationalized. The LMTF Phase III Implementation Working Group is currently exploring these terms and how they are defined by the education community and others involved in youth and workforce development.

Using Data to Improve Equitable Learning Outcomes

Because education is a universal aspiration and right, measures of access and learning at any level (global, national and subnational) should always pay attention to both aggregate measures of overall conditions (e.g., enrollment, achievement) and disparities in those measures among student subpopulations.

Disparities refer to different possible dimensions—such as sex, age, residence area (urban/rural); income level and socioeconomic status; poverty and extreme poverty status; mother tongue; ethnicity; regions within a country; citizenship status; disabilities; and emergency or catastrophe (natural or human-made) situations. Each country has the responsibility of identifying which dimensions are particularly relevant to its own conditions and design measurements that take these variables into account.

Measures of disparity should also be able capture change over time. Tracking progress in this context refers not only to changes in national aggregates but also to the reduction of disparities between population groups.

The prototype framework developed by the Measures and Methods Working Group proposed two levels of learning to be tracked at the global level—minimum and desirable. While the idea of multiple levels of learning was welcomed by consultees, they proposed measurement on a continuum of levels, or progress across a benchmarking scale rather than only at minimum and desirable levels. Participants were concerned that setting only minimum and desirable levels could send a message that some countries should only aspire to a minimum level of learning outcomes for their students while other countries should aspire to a higher level. This was seen as contradicting the LMTF's position that all children should have opportunities to learn across a broad framework of learning domains to prepare them for success in life.

Considering Multiple Methods

Large-scale measurement should not be developed for its own sake, but rather to make an impact on policy that ultimately leads to improvements in learning. Thus, the LMTF considers that learning measurement can help in different ways: (1) to identify and determine the magnitude of potential problems: (2) to track progress; (3) to inform interventions; and (4) to inform parents and the community in general about the status of education and thereby inform public debates. Every assessment effort is designed in a way that can give priority to some of these purposes over others. That is why no assessment is intrinsically better than another since that judgment needs to be made in light of the purpose it is expected to serve: A good assessment is one that is fit for the purpose.

The purpose of the assessment should dictate the technical attributes of the tools. For instance, if there is interest in tracking progress over time, measurement should be comparable for different moments in time; if there is interest in identifying disparities across subpopulations (for instance, by country, sex, age, ethnic group, linguistic group), measures should be disaggregated and equivalent across those groups; that is, they should measure the same constructs in equivalent ways for all groups. Moreover, measures should be able to detect significant differences—otherwise known as discriminant or sufficient inferential power—between marginalized groups to track equity in learning.

The Measures and Methods Working Group first examined the technical properties of assessments and then considered feasibility issues related to internationally comparable metrics, noncomparable metrics,

and additional efforts beyond school-based metrics as described below.

Technically Sound Measurement

Measurement should be done in a technically sound or robust manner. Weak data can be misleading and, therefore, may have a negative impact on policy. This does not mean that measurement efforts at early stages (when validity and reliability claims are not yet clear) should be discarded, but rather reinforces the fact that they need to be strengthened and the information they generate should be used with the utmost care.

The literature (cf. Postlehwaite 2004) and experience with learning measurement stress those robust learning measures whose basic attributes that can be summarized as:

- Clear definition of purpose;

- Clear identification of the target population;

- Clear and sound sampling strategy (if sample-based);

- Construct(s) are measured ensuring equivalence of information across populations (country, language groups and time periods—if identification of trends is sought);

- Validity (construct, content, predictive, etc.);

- Reliability (internal consistency of the results); and

- Documentation on all these matters that allows public scrutiny of methodological attributes.

If these measures are used in multiple countries (whether or not the results are internationally comparable), then it is important to additionally consider:

- Applicability in a diversity of contexts. This entails scales being sensitive enough to detect a wide range of skills for the trait being measured (for instance, reading skills from a very low to a very high level).

- Country coverage. A global measure should comprise at least a significant proportion of countries in each region, and a significant proportion of the overall population in each regional or country grouping. Groups at risk for social exclusion should be sampled to ensure adequate representation.

- Adaptability and/or co-development of measures. A measure should be adapted to each country in which it is used. Having a measure that has been co-developed through the collaboration of partners in different regions increases the likelihood of ensuring construct validity and relevance to diverse settings.

These criteria also suggest that measurement efforts have a cost and presuppose some institutional and professional conditions to be effectively conducted, such as: sufficient financial resources, adequate professional capacities, and political will to sustain the effort and technical independence of the body in charge of measurement. It is important for the LMTF to send a clear message that measurement, even when it looks costly, represents a small amount of resources as compared with the public resources that are wasted when policy is not grounded in sound evidence.[18] Research shows that the average costs of a large-scale educational assessment appear small relative to national educational budgets; while they are generally less than 1% of a national budget, they can be as high as 3% and as low as 0.3% (Wagner, Babson and Murphy 2011). Paraphrasing the statement on education that is usually attributed to Derek Bok, we could say: If

you think measurement is expensive, try using missing data as evidence. Having sound evidence helps make better use of the resources that are invested in education.

The international community and national governmental and nongovernmental actors must demonstrate a serious commitment to assessments beyond political statements. This is important to secure the requisite resources and ensure that the institutional conditions (e.g., the technical independence of those responsible for measuring) and professional capacities needed for sound measurement exist.

Internationally Comparable Metrics

To inform global debates, it is feasible to identify a set of metrics that would allow tracking progress at an internationally comparable scale. At the same time, this identification should not be seen as a constraint on the measurement of learning at the national or subnational level. The LMTF recommends that each country should establish or strengthen its capacities to systematically and periodically measure, in a technically sound manner, learning in the six areas of measurement described in this report. Moreover, international measurement efforts can be taken as opportunities for this national capacity development since countries can participate in international efforts not only to have comparable data, but also as one additional element in the consolidation of their measurement systems. There is no need to see national and international measures as opposed to each other; on the contrary, they can be seen as mutually reinforcing elements.

The consultation process revealed that the education field is divided over the importance of international comparability in tracking global learning levels.

One group of consultees reported,

"We were mixed on the importance of internationally comparable measures, with some team members feeling it is very important while others felt it is not important. Those who felt it is not important are skeptical of current results from assessments . . . because they feel that some countries are very selective about which students do/do not take the assessment in order to ensure their region is well-represented."

It is reasonable to expect that those areas that are considered appropriate for global monitoring are already the focus of different measurement endeavors that have been built on the basis of reflections that are not very distant from the debates of the LMTF. Nevertheless, the working group intentionally started from a broad perspective to identify areas where global measurement is desirable and might be feasible if more resources were devoted to measurement in this domain (or encompassing several domains).

Table 2 summarizes the relationship between available internationally comparable metrics and the domains of learning by educational stage.

The way the working group approached the task of defining measures was to claim that (1) all subdomains within each of the seven domains should be measured in early childhood, primary, and lower secondary school; thus, there are 21 areas where measurement efforts should be promoted; and (2) the previous statement does not mean that all domains should be simultaneously measured at the global and national levels.

This strategy also acknowledges that sound measures in each of the domains are not equally developed or currently available on an international scale. This translates into encouraging the development of measurement tools, while prioritizing the areas that should be developed first because they are considered more urgent. Moreover, encouraging measurement in the seven domains across the three education stages does not imply that national assessment systems of learning should be conducting measurements in all 21 areas simultaneously or with the same degree of depth. Strategies to rotate topics over a period of time can significantly enlarge the thematic coverage without creating an undue operational burden on the agents responsible for measurement. Finally, prioritization of measurement efforts is also expected, and thus each country could develop the measurement efforts according to its own priorities and circumstances.

The discussion above also affects the way measurement can be carried out, specifically the choice of tools to be used. The working group reflected on this matter and concluded that there is no reason to promote a single set of learning benchmarks to be universally achieved unless comparable information is used.

Comparable information refers to the ability of the metrics to generate evidence that is equivalent across populations (age, gender, residence, ethnic, socio-economic, linguistic background, country) and/or across time. Comparability can be desirable at both the national and international levels in order to enable tracking progress and identify equity gaps against a common metric. While "league tables" seem to be over-represented in the debates about comparability, comparable analysis makes it possible to learn from contrasting situations and identify equity gaps. When comparable information is required, technical mechanisms must be in place to ensure that a common metric with equivalent properties is used to measure across a variety of analysis units. This entails the identifica-

Table 2. Domains in Which Globally Comparable Measurement Is Both Desirable and Feasible			
	Early Childhood Level	**Primary Level**	**Postprimary Level**
Physical well-being	*EDI*		
Social and emotional	*EDI*		*ICCS*
Culture and the arts			
Literacy and communication	*EDI*	**PIRLS** **Pre-PIRLS** **(SACMEQ, PASEC, LLECE if pursue cooperation among them)**	**PISA** *ICILS*
Learning approaches and cognition	*EDI*		*TIMSS and PISA teacher and student questionnaires*
Numeracy and Mathematics	*EDI*	**TIMSS** **TIMSS-Numeracy** **(SACMEQ, PASEC, LLECE if pursue cooperation among them)**	**TIMSS, PISA**
Science and Technology		**TIMSS**	**TIMSS, PISA** *ICILS*

Note:

- **Bold:** robust internationally comparable measures exist and could be used at global level with expansion to more countries
  - PIRLS: Progress in International Reading Literacy Study
  - Pre-PIRLS: Less difficult version of PIRLS
  - TIMSS: Trends in International Mathematics and Science Study
  - TIMSS Numeracy: Less difficult version of TIMSS mathematics subscales
  - PISA: Program for International Student Assessment
  - PASEC: Programme d'analyse des systèmes éducatifs de la Conférence des ministres de l'Education des pays ayant le français en partage
  - SACMEQ: Southern and Eastern Africa Consortium for Monitoring Educational Quality
  - LLECE: Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación
- *Italics:* robust international measures exist but the country coverage is limited and adaptation is needed to make it feasible to measure at global level global on a common metric:
  - EDI: Early Development Instrument
  - ICCS: International Civic and Citizenship Education Study
  - ICILS: International Computer and Information Literacy Study (expected 2013)
- Empty cell: Local, national, or cross-national efforts exist but would require significant effort to scale up to internationally comparable levels

tion of: a common operationalization of the constructs; similar quality control mechanisms for all the phases of the data generation process; a careful review of the construct[19] equivalence across contexts (language or settings); and a way to monitor measurement equivalence.[20]

Noncomparable Metrics

When no common metric is sought, actors at the global level can provide technical guidance, prototypes and exemplars that countries can use to develop their own learning metrics.[21] Country-developed metrics have a huge potential for ensuring relevance (i.e., correspondence with the main issues that define their national policy agenda). At the same time, it should be clear that information generated using different metrics (even when inspired by a common set of guidelines) cannot be aggregated to create a global portrait of the situation. Pulling together information from different metrics can be seriously misleading.

While national data could not be aggregated at the global level without using a common metric, learning outcomes generated by national efforts could be presented at a global level as:

- Country profiles describing the assessments used; the decisionmaking process; learning trends (provided data are comparable across time); and policy changes informed by learning assessments.

- Equity-based comparisons of learners (e.g., disparities between boys and girls, urban and rural, wealth quintiles, etc.) using assessments and examinations meeting the corresponding technical criteria.

- A rubric or report card that assigns qualitative descriptors (e.g., initial stage, in developed, consolidated) to country-level systems based on multiple

factors. These could include the diversity of the domains assessed, periodicity, stability, and so on.

Beyond School-Based Metrics

There was much support from the task force, working groups and consultation participants to examine efforts to measure learning beyond assessments administered in schools. In exploring this topic, the Measures and Methods Working Group acknowledged a definition of learning that is present among some social scientists: "a relatively enduring change in behavior or knowledge as a result of past experience" (Hockenbury and Hockenbury 2010, 146). Beyond that, it can be viewed as a process whereby people develop a manifold set of types of knowledge, skills and attitudes that ultimately enrich their lives. As such, learning happens in a large diversity of settings and circumstances, and therefore it cannot be seen as something that is limited to the realm of academic or instrumental learning happening in institutions.[22]

Nevertheless, there is a substantive distinction between learning that is the result of an intentional action and learning in a nonsystematic or even incidental manner (Wagner, Murphy and De Korne 2012). While incidental learning is a critical form of learning for everyone, it is generally not considered a strong basis for higher-level cognitive processing because the assumptions, clarity and logic that underlie it are not systematically built. The LMTF is focusing on intentional learning because it is assumed that intentional learning, which is primarily the result of teaching, is necessary to build the cognitive structures required for higher-level cognitive processing. Public policy is concerned with intentional learning, since public action (by governmental or nongovernmental agents) is organized around those elements that societies deter-

mine to be matters of public concern. Education, as an organized activity aimed at ensuring that individuals (children, youth and adults alike) learn an identifiable set of types of knowledge, skills and attitudes that will enable them to function successfully in a civil society, is a matter of public interest, and thus it is organized through public policy mechanisms (e.g., an education system that includes national learning objectives, funding and managerial mechanisms).

Because universal schooling is not yet a reality for millions of children around the world, there was strong support from the task force, working group members, and public consultation for the LMTF to include recommendations for measuring the learning of children and youth who are not in formal education systems.

As one working group member noted,

"... if we are preparing an inclusive framework, we cannot forget or leave behind all those situations where the education system is not in place, or dormant and other mechanisms must take the lead to ensure the right to quality and relevant education to all children and youth."

Intentional learning typically occurs in institutional settings that are established to that end: educational programs. While these programs can be formal or nonformal in nature,[23] all of them are aimed at ensuring that learning is nationally agreed upon. That is why an increasing number of countries are developing mechanisms to allow the transition across these programs by recognizing and certifying the types of knowledge, skills and attitudes that are acquired in one setting as equivalent to those that are acquired in another setting. Recognizing the competencies of those who have

learned though nonformal educational programs, work and life experiences is a way of ensuring broader opportunities for all children and youth.

These considerations are important for the measurement of learning because they suggest that measurement should be linked to the specific purposes (and the corresponding content) that are embedded in the organization of intentional learning. These are the areas that are amenable to public action, and the convergence through the development of national curriculum and assessment frameworks (cf. Tuck 2007) suggests that the formal structure of education systems can be taken as the organizing principle of any measurement endeavor.

However, children and youth who are not in formal education programs should also be included in any global framework. This is one of the reasons why the work of the LMTF, while primarily giving attention to learning, is not exclusively focused on learning outcomes. As stated in the LMTF documentation, it is crucial to promote a policy agenda focused on access plus learning. This means that the pending challenges regarding access should not be overlooked, not only because of their impact on the educational situations of millions of individuals but also because the information produced by learning assessments that intend to portray a national situation should be read taking into account this element of the educational reality.

Depending on the coverage of the country's education system, measuring learning for out-of-school children and youth may yield useful information. Since many out-of-school children live in poverty, conflict/disaster areas and other hard-to-reach areas, they face daily challenges that require them to solve problems, make critical decisions, be persistent and creative—and

many times these skills are crucial for their survival. They also often need to have leadership skills and be aware of their environments, knowledgeable of their local culture and customs and cognizant of the dangers that may surround them. Some countries (and even regions) might benefit from measurement not only of literacy and communication but also within the domains of their out-of-school children and youth vis-à-vis social and emotional experiences, learning and cognition skills, science and technology, and culture and the arts.

A number of children are excluded from mainstream education due to disabilities and other health-related issues. Similarly to the modules related to health that are complementary to learning assessments (e.g., SACMEQ's module on HIV) and measurements related to knowledge of illness and disease prevention can also be measured for out-of-school children though household surveys and/or assessments. As learning takes place both in and outside formal school settings, household-based assessments with sound sampling and testing procedures would ensure that learning levels of out-of-school children are being captured and properly measured. Examples of assessments which include out-of-school children and youth include:

- The Citizen-Led Assessments, described in the previous sections on measuring reading and numeracy, are conducted in the home and thus generate information on learning levels for children in school, out of school, and in nonformal education settings.

- Literacy experts at Education Development Center (EDC), Inc. have developed the electronic Out-of-School Youth Literacy Assessment (OLA) to measure literacy skills of older youth and young adults, particularly those who are living in extreme poverty or postconflict environments. OLA builds on reading research and best practices in evidence-based adult literacy instruction and assessment.[24] The instrument assesses some aspects of phonics, word rec-

ognition and decoding, fluency and comprehension, but also includes a real-life/functional reading section. The tool has been piloted in Liberia in English and in Rwanda in Kinyarwanda using electronic platforms, and will be piloted in Mali in Bambara in early 2013. OLA is currently being used as a baseline and post-test assessment to measure change in literacy acquisition and development. Because it assesses critical pre-reading skills, OLA results can also be used to inform curriculum and teaching approaches, and real time data are used to improve the implementation of assessment activities.

Assessments of learning outcomes that are administered as household surveys would be particularly useful in countries where large numbers of children and youth are excluded from education, including countries where:

- There is great variety in school provision (e.g. a significant percentage of children are enrolled in private schools with little regulation);

- Attendance rates of teachers and students are variable;

- Primary and lower secondary enrollment and retention are low;

- Lower secondary schooling is not compulsory; and/or

- There are conflict- and disaster-prone areas.

Household-based assessments can be designed to (1) help national governments and non-governmental actors develop learning programs for all children; (2) draw attention to the gap between the skills that youth have and the skills they need for employment; and (3) provide evidence to support advocacy efforts by students, parents and civil society.

Rather than advocating or not for household-based assessments, the task force can develop a set of criteria that countries can use to assess the need and benefit

of measuring both their schooled and out-of-school children in one or more domains. In that way, countries would be able to, in relation to out-of-school children, determine (1) if measurement is feasible; (2) if it would be useful; and (3) on which domains to focus measurement. The Implementation Working Group will examine countries in which household-based assessment of learning are used and make further recommendations on this topic.

Assessment as a Global Public Good

There was considerable debate among the task force and working group members about how data are produced and managed. While education statistics systems and national and international statistics are public goods (i.e., funded with public resources to serve a public purpose), this is not always the case for learning assessments. The LMTF decided that it could not propose a global learning goal that requires countries to buy into only one brand of assessment.

This situation suggests that for assessment data to become a public good, some basic elements should be advanced:

- Full documentation of the studies that are funded with public resources should be publicly available. This documentation includes the data sets, instruments and procedures used to generate the data. Procedures should be fully replicable by independent researchers, and the data should be publicly available with the proper protections of the identities of the individuals who consented to participate in the studies. The only exception to this principle applies to the safeguarding of instruments' integrity; in this case, instruments can be kept in reserve and only exemplars be publicly released.

- Informed and explicit consent by participants in the studies should be properly guaranteed.

- The body in charge of conducting the studies must have technical independence so as to be able to make independent decisions on what is publishable and what is not.

- Collaboration among different agencies should be promoted as a way of ensuring that a diversity of in-

terests, perspectives and needs is embedded in the development of the studies from the outset. Off-the-shelf approaches or the contracting out of private business that keep proprietary rights to essential components of the assessment are legitimate options that, however, would not necessarily guarantee the public nature of the data.

- Collaboration among public and private assessment agents can take different forms, ranging from the co-development of a given study (as is the case in IEA´s studies) to agreement on technical procedures that would make one study comparable to another (as in the proposed collaboration between the regional assessments and PIRLS). This collaboration can also include significant components of technical exchanges, including the sharing of experiences, instruments and procedures.

In short, the bottom line is to ensure that data, data generation and techniques are commensurate with public needs and evolve with them. This entails preventing the development of monopolies led by private interests. Additionally, when international comparability is sought, several elements should be standardized and verified to ensure that: the instruments measure the same constructs in an equivalent way across different settings; the different stages of the study are conducted with an acceptable level of quality; and the analysis is performed following equivalent procedures. This situation makes comparability much more complex than following a common blueprint, since the way it is followed affects the final results. These considerations are also important because international comparability cannot be achieved without paying due attention to diversity. Measuring the same constructs

in an equivalent way is not equal to measurement based on the use of identical tools, for example, since the meaning of a particular item may change due to the properties of the language or the language usage in a particular context. Careful adaptation is required, and the adaptation should also be monitored to ensure that it leads to equivalent measures (Hambleton 2005).

Finally, whatever scheme is used at the global level should be built upon a clear understanding of its potential and limitations. There is tremendous potential in global monitoring and comparative analysis of learning, but the results must be analyzed in a way that takes into account contextual information. At the same time, global information systems can play an important role in strengthening the national information systems, and country participation in international efforts should always be aimed at making the most of it in relation to their own development. In short, international assessments are not an alternative to national or subnational measures, but a supplement that, ultimately, should reinforce them.

Global Development and Education Goals

Setting goals is of paramount importance to enable the identification of measurement tools and procedures. The objectives of the LMTF project include ensuring that learning is included in the post-2015 development framework and proposing indicators that can be used to track progress toward the global goals to be identified through the post-2015 MDG and EFA processes. The Phase II prototype document for public consultation proposed global goals for the early childhood, primary and lower secondary levels, but consultation feedback on the proposed goals was mixed; some groups said they were appropriate, while others thought the specific goals were too narrow and focused on schooling.

The consensus within the education community is that there should be a shift in focus from access to access *plus* learning. For instance, the *World We Want* education consultation—co-led by UNICEF and UNESCO with the support of the governments of Senegal, Canada and Germany, and the Hewlett Foundation—recently released a summary report putting forth the proposed goal of *equitable, good-quality, lifelong education and learning for all*, which has emerged from the overall consultation process on how best to reflect education in the post-2015 development agenda.[25] The synthesis report states that "there is a sense of unanimity that regardless of the structure the post-2015 agenda may take, education must claim an explicit goal focusing on equity, access, and quality learning" (UNESCO and UNICEF 2013, 1).

Moving beyond this broad goal, the report notes that there have been a number of propositions for more specific goals by different education sector constituencies focusing mainly, but not solely, on (1) expanded access and completion; (2) the provision of quality education and learning; (3) the enabling conditions necessary for quality education and learning; (4) a renewed focus on gender equality; and (5) the need for measurable targets and indicators, which allow for regional and national adaptation. On the final factor, the report emphasizes that the agenda must strike a balance between providing internationally comparable goals while allowing for national ownership by setting ambitious goals, targets and indicators that are applicable to the national context and challenges.

Reflecting the work and recommendations of the LMTF, the synthesis report highlights the need for the education community to remain focused on the actual learning outcomes and the content and relevance of education. The report notes that this implies "focusing on relevant, measurable and equitable learning outcomes, such as foundational literacy and numeracy, as well other relevant social, civic, economic, environmental and health-related skills and competencies. The focus on learning must be combined with one on equity, with particular attention to rising inequality within countries" (4).

The *World We Want* education consultation executive summary, once completed with inputs from the

Advisory Group and the broader community, will feed into the final report from the United Nations Development Program summarizing the outcomes of the various thematic consultations.

Another education goal has been proposed by the UN Sustainable Development Solutions Network (SDSN) in the draft report, *An Action Agenda for Sustainable Development*. The report proposes the following goal to *ensure effective learning for all children and youth for life and livelihood:*

> *All children participate in adequate early childhood development programs, and receive primary and secondary education to prepare them for the challenges of modern life and decent livelihoods. All youth and adults have access to continuous lifelong learning to acquire functional literacy, numeracy and skills to earn a living through decent employment or self-employment.*

The SDSN report recommends that societies take a "life-cycle" perspective to education, focusing on the needs of individuals at each stage of the life cycle. As such, targets might include:

- All girls and boys have access to early childhood development (ECD) programs.

- All girls and boys receive quality primary and secondary education that focuses on learning outcomes defined for each stage of schooling, and on reducing the dropout rate for each grade to zero.

- Youth unemployment rate is below 10%.

The SDSN report will feed into the UN Secretary-General's report on the Sustainable Development Goals later this year.

Much of the goal-setting at the global level has focused on reading. In addition to the UN Secretary-General's Global Education First Initiative, increasingly donor and civil society organizations are setting specific, numerical and time-bound goals for improving reading skills in developing countries. AusAID's comprehensive policy aid framework indicates that 25% of aid will be spent on education assistance to get 4 more million children in school and provide 20 million children with a better-quality education by 2015–16 (AusAID 2012). The Global Partnership for Education (GPE) aims to improve reading levels of children in the early grades in at least 20 GPE countries within five years, with the specific target of cutting in half the number of children that are not reading by grade 3 (GPE 2012). USAID seeks to achieve the target of improving foundational reading skills for 100 million children in primary grades in its partner countries by 2015 (USAID 2012). The nongovernmental organization (NGO) Save the Children aims to design and test literacy programs that measurably improve children's learning outcomes and implement them in at least 20 countries, which is intended to reach 600,000 children by 2015. Save the Children has also set the target to assess reading skills of 100,000 children (Save the Children 2012). While these goals are intended to positively influence the number of children reading worldwide, some caution that overambitious goals may deflate countries' motivation for improvement (Beatty and Pritchett 2012).

Assuming learning will be included in the next set of global development goals, the task force offers a menu of indicators with recommendations as to which ones are technically feasible to operate at a global level (i.e., are appropriate for both post-2015 and EFA goals). The work of the LMTF can offer governments a framework for what a goal could look like and provide a shared vision toward a learning goal based on the seven domains. Rather than put forth a specific goal,

the task force has decided instead to develop recommendations for measuring learning that are flexible and adaptive for inclusion in post-2015 discussions, acknowledging that the specific wording will ultimately be developed by UN Member States and Secretary-General's office. These indicators will be initially presented in September 2013 and refined over time.

While the LMTF did not agree on a set of goals, it does advance the following overarching vision, based on the seven domains of learning:

Every child and youth should fully develop and acquire basic foundational learning through the provision of care and educational services that pay due attention to their physical well-being and their learning in a broad range of areas that include social and emotional, culture and the arts, literacy and communication, numeracy and mathematics, science and technology, as well as learning approaches and cognition.

Expanding National Capacity and Improving Decisionmaking

The LMTF acknowledged the need to improve national information systems for learning and their capacity to inform policy and leverage change. While the globally tracked indicators described in the previous sections are an important component of the LMTF's recommendations, the task force and working group viewed them as one piece in the overall effort to have sound and relevant data on access plus learning.

Many efforts are under way to assess learning at the national, regional and international levels; more than 150 countries currently measure learning levels through national assessments and examinations or participate in international, regional or cross-national assessment initiatives (UNESCO/UIS 2012). Many of these efforts are bottom-up approaches—for example, the Annual Status of Education Report (ASER) in India began as a small-scale assessment by Pratham to better understand the impact of the NGOs' literacy work (ASER Centre 2012). The demand for information on learning levels grew at the local levels, and eventually became a useful tool for raising awareness of low learning levels and advocating for better education systems. Other researchers in Pakistan, Kenya, Tanzania, Uganda, Mali and Senegal became interested in Pratham's model and now conduct similar assessments that rely on volunteers to reach the remote areas of these countries. While its impact on public debates is clearly important, further research on the robustness of the instruments in all countries would be useful to expand this model (cf. Vagh 2012). Other practitioners—often from NGOs or research firms, such as Save the Children and Research Triangle International—have developed measurement approaches, particularly for the primary early grades, that have some features in common across countries but are explicitly not aimed at producing international comparisons on an equivalent metric because of the specific characteristics of the settings and procedures used as well as the languages and scripts.

This type of approach plays a significant role in promoting and supporting the development of national assessment systems. It is also a type of approach that could support measurement domains and subdomains that would not be appropriate for international comparability. At the same time, in some domains, there is regional or national experience in measuring learning that with further development could lead to internationally comparable measures. A particular domain or subdomain may not be measured in a cross-national comparable way, for several reasons:

- There are robust measures, but the results are not suitable for international comparison (many assessments of foundational skills, such as the Early Grade Reading Assessment, EGRA, or the in-development Early Grade Writing Assessment, EGWA, fall into this category);

- There may be some domains for which it is undesirable or impossible to develop internationally comparable assessments (e.g., measures of religious knowledge and values; creativity and experiential learning in the domain of arts and culture); and

- There are measures that show promise for international comparability but there is no demand from countries to participate in an internationally comparable assessment on this domain.

Supporting country-specific processes aims at creating opportunities for innovation in learning assessment with resources from the international community. This idea was particularly supported by those wanting to see learning measured in the area of culture and the arts, for which no internationally comparable measures exist. The working group expected that, under some circumstances, countries might improve their national-level practices as a step along the way to participating in internationally comparable assessments. It is important to note that while some of the assessments mentioned above are not intended for international comparisons on equivalent metrics, the assessment organizations can still learn from each other and develop good practices and this can and should be done with more guidance from official bodies. Thus, in addition, the international community and the international bodies mandated to work on these issues can greatly support the exchange of information about these national or nationally specific approaches.

Identifying Country-Level Communities of Practice

A country-level community of practice (CoP) body would comprise: education ministry officials; national assessment experts; civil society groups; and organizations representing teachers, parents, students and other key stakeholders. In some countries, there is already a national body proposed or fulfilling this role consisting of the education ministry, academia, civil society, teachers, donors and other stakeholders that makes decisions on assessment. Examples include:

Kenya: The National Assessment Center (NAC) is housed within the Kenya National Examination Council. The NAC is charged with coordinating assessments within the country. In order to include nongovernmental actors and leverage the policy impact of the

NAC, an action group was proposed in 2012 made up of various government departments and agencies and NGOs to address issues of learning outcomes through 2015 and beyond. This group includes representatives from Well Told Story, Women Educational Researchers of Kenya, Elimu Yetu Coalition (the umbrella organization for all civil societies working in education of which WERK is the Chair), Aga Khan Foundation, National Taxpayers Association, University of Nairobi, Uwezo East Africa, Africa Population and Health Research Center, and Research Triangle International. Its objective can be summarized as "Making Learning the Measure of Basic Education: Consolidating Efforts for Learning as Countdown to 2015 in Kenya."

Zambia: The Ministry of Education Science Vocational Training and Early Education has administered the Zambia National Assessment survey biennially at grade 5 since 1999 to assess learning achievement at the middle basic level. The survey is guided by the National Steering Committee comprising academia from the University of Zambia, Examinations Councils of Zambia staff, education sector donors, civil society representatives and Ministry of Education officials. It is chaired by the Ministry of Education's permanent secretary. The steering committee provides oversight and gives policy direction of the survey and is supported by a Technical Committee of assessment experts in the execution of its mandate.

The national CoP would engage in a process to: review domains of learning; recommend national priorities for which domains to measure and when; examine technical properties of measures currently used; and recommend which measures to use. The country could work with a global organization or technical assistance provider to conduct a self-assessment. This type of self-assessment and policy guidance is currently being implemented in more than 30 countries

through the World Bank's Systems Approach for Better Educational Results (SABER)—Student Assessment program and READ Trust Fund program (Clarke 2012; World Bank 2011).

The scope of this approach could be much broader than system-level assessments of learning. The CoP, regional organizations and a neutral international body could also be mobilized to improve the classroom assessments, national examinations and assessments used for evaluations of specific programs that are not administered to a nationally representative sample.

Support from the International Community

To support country CoPs and track progress at the global level, the working group proposed a neutral, multi-stakeholder, international advisory group that would connect countries with resources and technical assistance to decide which domains to measure, how to measure them, and how to implement the assessments. The international body would coordinate with regional educational organizations such as the Arab League of Educational, Cultural, and Scientific Organization (ALECSO), Association for Education Development in Africa (ADEA), Organización de Estados Iberoamericanos (OEI), South Asian Association for Regional Cooperation (SAARC) and Southeast Asian Minister of Education Organization (SEAMEO) and, as appropriate, provide resources for these organizations to support learning measurement at the country level. There have been some tentative steps taken in this direction as described below. However, some of these initiatives lack sufficient funding and may benefit from coordinating efforts at the global level.

This body would serve in a coordinating and logistical role. The body would be independent of existing assessment bodies and private corporations. This group's responsibilities might include:

- Convening partners: Convene key actors including teachers organizations, global and regional organizations, assessment experts, private corporations, and the like, to provide technical expertise and financial support to countries for measuring learning.

- "Center of excellence": House or support a global clearinghouse for best practices and research; be a repository for lessons learned and good practices.

- Policy and advocacy: Mobilize governments and the international community to devote resources to measuring and equitably improving learning outcomes.

- Global learning metrics: Coordinate the development of common metrics for learning indicators and promote their use.

- Standards and technical criteria: Set standards for the design and administration of learning assessments; promoting and supporting quality standards for international, regional, and national assessments

- Contribute to tracking progress: Work with existing agencies (UIS/GMR) as they work to compile and report out on global education data.

- Capacity building: Support CoPs at the national level to build capacity and develop actionable plans for measuring and improving learning; support regional education assessment organizations (e.g., SACMEQ, PASEC, LLECE).

- Participatory process: Facilitate a participatory process so that all interested actors have a voice in determining and implementing global metrics.

- Official mandate: Have a recognized mandate among stakeholders.

Some components of this international body already exist or are planned (see table 3); for example, the Student Assessment section of the World Bank's SABER, the UIS's Observatory of Learning Outcomes, and the IIEP's Portal on Planning for Improved Learning. GPE is also working with a variety of official agencies, including the UIS, to promote the sorts of exchanges noted here but, as mentioned above, the degree of mandate and funding that exists for these initial steps is smaller than would be required to have a major global impact. This body would not seek to duplicate these efforts but rather to link them together.

Table 3. Examples of Current Activities	
Organization	**Activities**
Global Partnership for Education	Working with UIS, UNESCO, IEA, regional assessments and other agencies to promote exchanges of information on learning outcomes.
UIS/GMR	Global education data gathering and reporting.
World Bank	Providing technical assistance to countries for reviewing and improving their learning assessment systems through the SABER and READ Trust Fund programs.
UIS Observatory of Learning Outcomes	Gathering information on all learning assessments at the country level (including national assessments and examinations).
IIEP Portal	Gathering information to guide education ministries on collecting and using learning assessment data.
International Association for the Evaluation of Educational Achievement (IEA)	Learning assessment studies in reading (PIRLS and PrePIRLS), mathematics and science (TIMSS), civics and citizenship (ICCS), and an upcoming assessment on computers and information literacy (ICILS). Plans for TIMSS-Numeracy, a less-difficult version of TIMSS, are underway for administration in 2016.
OECD	Learning assessment studies in reading, mathematics, and science (PISA) in addition to financial literacy and collaborative problem-solving. Assessments include contextual questionnaires related to learning environments and non-cognitive outcomes. A PISA for Development initiative is being implemented that aims to enhance the relevance of PISA for developing countries and thereby increase the participation of these in future studies. PISA assessments are competency-, skills- and content-based.
Regional assessment consortia (PASEC, SACMEQ, LLECE)	Develop and administer regionally comparable assessments based on national curricula.
Organisation of Eastern Caribbean States	Convened country stakeholders to develop education sector strategy, "Every Learner Succeeds," which includes agreed-upon learning outcomes for early childhood, primary, and lower secondary.
Southeast Asia and Pacific Region (SEAMEO, UNESCO, UNICEF and partners)	Early childhood: UNICEF Regional Office for East Asia and the Pacific (EAPRO), UNESCO, the Asia-Pacific Regional Network for Early Childhood (ARNEC) and the University of Hong Kong have developed the East Asia & Pacific Early Child Development Scales for children age three to five which are currently in the validation phase. Primary, there is an initiative under way between UNICEF and SEAMEO to develop metrics for the primary level for SEAMEO member countries, in which UNESCO may engage as well. Additionally, UNESCO is now working to set up a regional network for the monitoring of educational quality in the region which will focus on information exchange, research and potentially capacity building around assessment issues.

Additional Overarching Considerations

During the final phase of the LMTF, the task force and working group will examine strategies to implement the measurement of learning in a way that ultimately results in better learning outcomes. This section describes considerations for the final working group on implementation.

Measuring to Improve Learning at Different Levels of Agency

Many agents—from the classroom to global agencies—play a role in a child's learning process. Teachers play a pivotal role in the organization of the learning experiences, and so their professional practice (which includes assessment) affects learning on a daily basis. Assessment practices conducted by teachers are of paramount importance; they are part of multiple strategies teachers use to organize, lead and facilitate learning experiences.

Teacher-led assessments provide a fundamental connection with the actual learning experiences. These assessments are the ones that matter the most for each individual learner. Aggregate measures of learning can be of help, but they cannot replace or play a role that is equivalent to the assessment practices that good teachers can design and carry out.

Educational management and educational policies can depart from the bureaucratic trend to prescribe what agents should do and embrace a more fruitful approach where agents operating at the subnational,

national or global level could enable those who are engaged in the areas where learning takes place. This idea can be represented as shown in figure 4.

As shown in figure 4, it is important to understand that the impact on classroom dynamics decreases the further away a decision is made. However, since power asymmetries exist, decisions made at different levels can have significant effects on the school and classroom settings, even if they are not so prominently present in everyday dynamics.

By the same token, it is important to take into account that whatever can be said in relation to learning at the global level, it is usually a fraction of what can be depicted at the national level—and so forth. The richness of classroom dynamics is portrayed by the different techniques and resources that are mobilized by those who act at that level (teachers and their professional resources, the actual curriculum, their assessment practices, etc.), while less detailed information is usually presented at a broader aggregation level. Thus, global measures cannot intend to be as comprehensive as teacher-managed information, and global efforts should be built upon the strengthening of national, subnational and ultimately classroom assessments and learning.

As previously stated, measurement can serve a variety of purposes, and not all of them can inform interventions aimed at improving learning in the same way.

Figure 4. Improving Learning and Different Levels of Agency

Classroom:
Student-Teacher,
Student-Students,
Students-Teacher,
interactions

Learning

Enabling action — Global Agency

Enabling action — National Agency

Enabling action — Subnational Agency

Enabling action — School Agency
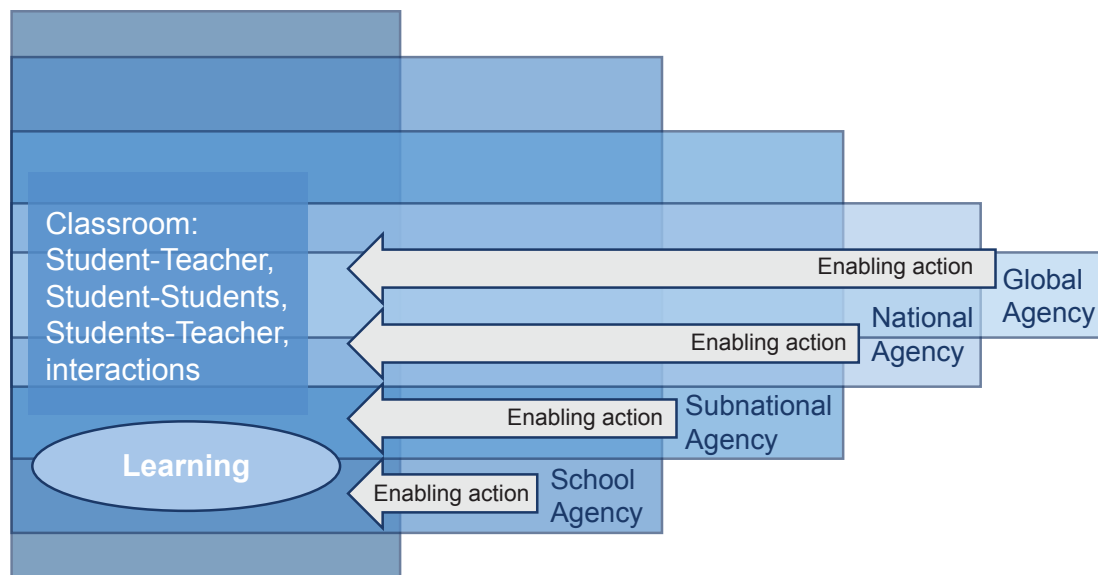
Examining How Assessment Data Are Used

As has been shown elsewhere (Kellaghan, Greaney and Murray 2009), different factors affect the way assessment data are (or are not) used (Kellaghan, Greaney and Murray 2009, chap. 1), and there are different avenues to explore in order to translate data into policy (chap. 4); management (chap. 5), teaching (chap. 6) and public awareness (chap, 6). Nevertheless, all these options rely on the extent to which an assessment was designed having those potential uses in mind, that is, the extent to which an assessment effort corresponds to its intended usage.

Tracking progress helps in identifying challenges and potential areas that are lagging behind. When done at the global level, it provides important information to the global education community to identify priorities. Moreover, at the national level, national authorities can rely on this kind of information to identify geographical

or content areas that are lagging behind and require additional support. Nevertheless, information that is needed to track progress does not yield direct clues on what an intervention could look like, unless it is coupled with other analytical elements aimed at connecting the performance information with school and other relevant education settings as well as context-related phenomena.

At the same time, tracking progress at the global or national level, while important to identify global or national priorities, might be too distant from the actual local contexts where learning happens. Thus, the measurement of performance requires a significant amount of contextual information that, in many cases, can only be produced at the local level. Beyond the relevance for policy, aggregated measures of student achievement are intended to inform agents who operate at the aggregate level, and these agents need to be able

to translate that information into evidence that would enable local agents (teachers, head teachers, district officials) to perform their duties in a better way, and to discover actions needed to provide the necessary professional support.

The capacity of large-scale assessments to inform interventions is contingent upon their scope; when learning achievement data are generated concurrently with information about the learning process, there is an increased potential for the practical use of this information to improve learning.[26] That is why taking into account learning trajectories and learning opportunities together with performance data can meaningfully inform policy and interventions, and that is why some assessments are built around school trajectories and experiences and, therefore, are conducted in school settings. Consequently, these assessments focus exclusively on learners who happen to be students, but there is no other way of connecting their performance to actionable characteristics of the education system. This does not mean that household-based endeavors are less important—only that they provide a different set of information that is more suitable for other purposes, like generating parent and community involvement or interventions on variables that primarily belong to household characteristics (like nutrition in early childhood) than to educational settings.

These reflections highlight that adequate bridges between assessment of learning and improvement of learning need to be carefully built. Developing standard "recipes" (which supposedly work at any aggregation level) out of standardized information is an approach to be avoided. At the same time, the LMTF, and the international educational community in general, could urge researchers to more precisely state how one can avoid recipes and yet take advantage of the information provided by international (or regional)

assessments, or indeed any assessment from outside the school.

The working group recommends that the Implementation Working Group in the final phase pay particular attention to the usage of the measurement information based on the principle that measurement design should be commensurate with the intended uses.

Balancing Global and National Approaches

Monitoring the effectiveness of the education system requires information that is connected to that system. Countries have varying levels of capacity for providing education and measuring learning, and this must be taken into account when developing recommendations. Student achievement information yields evidence on the extent to which the education system is able to ensure learning (the focus is not to see if students "fail" but if the system fails), and this is also related to access levels. For instance, there is a universal consensus (e.g., MDG Goal 2) on the need to ensure that every child completes a full cycle of primary education and achieves the learning that this educational level is intended to provide; thus, completion levels and learning levels should be measured and interpreted concurrently. At the same time, it is essential for countries to ensure that those who could not access and/or complete primary education can acquire equivalent learning in other settings and have that learning properly recognized. It is also important that different definitions and expectations about what learning primary education is intended to provide can affect how that progress in ensuring learning is perceived.

Moreover, a focus on specific population cohorts (and not only students) leads naturally to consider non-school-based approaches to data generation

in some countries. Nevertheless, in contexts where those specific cohorts are universally schooled, it might be a sensible operational approach to structure the data generation efforts around schools. Focus on population cohorts also requires paying attention to the quality of existing administrative records of births, since they could affect the robustness of age-specific information. While household surveys have been addressing this issue for a significant period of time, a lack of sound birth records still affects enrollment data in several countries, and the need to pay attention to it should not be overlooked.

Another element pertains to the tension between nationally or globally agreed-upon curricular goals and local knowledge. The development of quality education systems entails paying due attention and valuing local knowledge; in particular, its component related to technology and sustainable development should not be overlooked. Fostering a dialogue about universal values and expectations, national policies and local perspectives is key to bring about a more tolerant and respectful world. In this regard, the potentially unintended effects of implementing standardized measures should be carefully addressed, and will be a focus of the Implementation Working Group.

Conclusion

Measuring learning outcomes is a crucial step in ensuring that every child, everywhere, is able to realize the right to a good-quality education in order to enable them to become a productive global citizen. The education sector has successfully rallied to get millions of more children into school and bring awareness to the fact that education means learning, not just schooling. However, there is still much work to be done in order for education stakeholders to adequately measure and track success at the global level.

This report provides a global vision for how learning should be measured around the world. Education is the responsibility of everyone in society, not just the government, so the measurement of learning outcomes must be led and supported by all individuals who have a stake in ensuring that all children learn: students, teachers, parents, civil society, employers and a variety of other actors, in addition to the government. In some cases, this begins with a shift in the national dialogue on education from access to access plus learning. In other cases, where learning outcomes are already being measured and discussed in the public forum, the national dialogue may shift from learning to learning outcomes that are relevant for a globalized economy. Regardless of the "culture of evaluation" surrounding learning in a given country, there are steps that the education stakeholders can take to improve the measurement of learning outcomes and ultimately improve learning levels.

The final phase of the LMTF will address this crucial question: How can measuring learning be translated into policies and practices that result in improvements in learning? The Implementation Working Group will take the global recommendations and apply them to their own country contexts and those in which they work to explore the feasibility of tracking learning globally in these areas and determine the resources required, from both within and outside the country, to implement robust systems of learning assessments that are globally informed yet nationally relevant.

The task force met in Bellagio, Italy in July 2013 to deliberate on these issues. It will release a final report in November 2013, which will outline recommendations for national and international actors.

Annex A: Individuals Contributing to the Phase II Public Consultation Period

The following is a list of individuals who contributed to the Phase II consultation period and provided their name and affiliation to the LMTF Secretariat. This list does not include the task force members, working group members and secretariat, who also provided feedback during the consultation period. A full list of task force, Secretariat and working group members is provided near the start of this report. Note that individuals who participated in an in-person consultation and did not provide their name and affiliation are also counted toward the total estimate of 600 participants. The task force extended the work of the Measures and Methods Group to Phase III and formed the Existing Measures Subgroup. Members of this subgroup provided valuable contributions to the six areas for measurement in this report. Their names are listed after the following list of Phase II contributors.

Name	Organization	Location
Mohammad Chaib Diraa	Ministry of Education	Algeria
Ara Edoyan	Ministry of Education, Republic of Armenia	Armenia
Robert Stepanyan	Ministry of Education, Republic of Armenia	Armenia
Shahin Islam	Save the Children	Bangladesh
Manzoor Ahmed	Campaign for Popular Education (CAMPE)	Bangladesh
Mahmuda Akhter	Institute of Child and Human Development	Bangladesh
Lyudmila Khudenko	National education institute, Republic of Belarus	Belarus
Ket Chant	Plan International	Cambodia
Sheldon Schaeffer	ECCD	Canada
Vilelmini Tsagkaraki	McGill University	Canada
Chizoba Imoka	Ontario Institute for Studies in Education (OISE)	Canada
Cristina Jaimungal	Ontario Institute for Studies in Education (OISE)	Canada
J. Bhattacharya	Ontario Institute for Studies in Education (OISE)	Canada
Josh Jasodhara	Ontario Institute for Studies in Education (OISE)	Canada
Qin Liu	Ontario Institute for Studies in Education (OISE)	Canada
Eric Sossouhounto	Plan International	Canada
Jennifer Budden	Plan International	Canada
Manuel Cardoso	UNESCO Institute for Statistics	Canada
Nancy Del Col	World Vision Canada, Ed.	Canada
Mortiz Bilagher	UNESCO Santiago	Chile
Albert (Peter) C. Petite Jr.	Beijing No. 2 Middle School, International Division	China
Dr. Alfredo Sarmiento	Escuela Nueva	Colombia
Dr. Carmen Emilia Perez	Escuela Nueva	Colombia
Juliana Vernaza	Escuela Nueva	Colombia
Margarita Peña	ICFES	Colombia
Horacio Alvarez Marinelli	Inter-American Development Bank	Dominican Republic
Natacha Hoheb	Blossom Centro Familiar	Ecuador
Angel Camilo Marquez	Blue Hill College	Ecuador
Sofía Alvarez	Centro Educativo Logos Academy	Ecuador
Marjorie Mujica	Colegio Americano de Guayaquil	Ecuador

Name	Organization	Location
Andrea Bejarano	Independent Consultant	Ecuador
Fátima Sierra	Learning Projects	Ecuador
Jacqueline Ibarra	Logos Academy	Ecuador
María Grazia Guevara	Logos Academy	Ecuador
Paola Gaibor	MEC – Coordinación Educación Inicial	Ecuador
Perla Adrian	MEC - SEDG – Zona 8	Ecuador
Dialys Vega	Ministerio de Educación MEC	Ecuador
Marisol Quinche	Ministerio de Educación MEC	Ecuador
Mariuxi Valdivieso	Ministerio de Educación MEC	Ecuador
Marlene Quinto	Ministerio de Educación MEC	Ecuador
Wendy Lozano	Ministerio de Educación MEC	Ecuador
Yanka Hualpa	Ministerio de Educación MEC	Ecuador
Lily Arenas de Cajas	Unidad Educativa Steiner	Ecuador
Lorena Durán	Universidad Casa Grande	Ecuador
Lucila Pérez	Universidad Casa Grande	Ecuador
Marcela Frugone	Universidad Casa Grande	Ecuador
Dolores Zambrano	Universidad Casa Grande, Unidad Educativa COPOL	Ecuador
Guillermo García	Universidad Católica Santiago de Guayaquil	Ecuador
Moustafa Emam Hemedi	American University in Cairo	Egypt
Samir Shafik Habib	Research Triangle Institute	Egypt
Salomon Cruz	Plan International	El Salvador
Wendy Richorloff	UNICEF	Fiji
Heikki Lyytinen	University of Jyväskylä, Department of Psychology	Finland
Alejandro Gomez Palma	OECD	France
Felix Zimmermann	OECD	France
Michael Davidson	OECD	France
Michael Ward	OECD	France
Miho Taguma	OECD	France
Borhene Chakroun	UNESCO	France
Jun Morohashi	UNESCO	France
Mari Yasunaga	UNESCO	France
Julien Yombouno	Plan International	Guinea
Bidisha Mukherjee	N/A	India
Bratin Chattopadhyay	N/A	India
Poorva Chadha	N/A	India
Suman Sachdeva	Care India	India
Binita Baishya	DIET, Kamrup (Mirza)	India
Saradi Bhattachariya	DIET, Kamrup (Mirza)	India
A.K. Choudhury	DIET, Morigaon	India
Lipika Medhi	DIET, Morigaon	India
Samhita Malini Baruah	DIET, Morigaon	India
Sebika Bora	DIET, Morigaon	India
Nilima Sarma Mahanta	District Institute of Education & Training (DIET)	India
Dipti Lal	Education Initiatives	India
Lant Pritchett	Harvard Kennedy School	India
Srikantan Moorthy	Infosys ltd.	India
Bijli Mallik	IPER	India
Bandana Dutta Gupta	Knowledge Hub, Loreto Sealdah	India
Noor Ashfaque	Knowledge Hub, Loreto Sealdah	India
Anjlee Prakash	Learning Links Foundation	India

Name	Organization	Location
Animesh Chatterjee	Pratham Education Forum	India
Buddha Baruah	SCERT, Assam	India
Geeta Dutta Borthakur	SCERT, Assam	India
Sushmita Sutradhar Das	SCERT, Assam	India
Meenakshi Atal	The Heritage School	India
Amrita Sengupta	UNICEF, West Bengal	India
Kanu Priya Jhunjhunwala	UNICEF, West Bengal	India
ShubhraChatterji	Vikramshila	India
Aliah B.P. Hasan	N/A	Indonesia
Emilia Nazir	N/A	Indonesia
Eny Suwarni	N/A	Indonesia
Sri Murdanigth	N/A	Indonesia
Y. Novena Srikandewi	N/A	Indonesia
M. Zahain	Frontier for Health	Indonesia
Carmelita Cajumban	International Dutch School - Trilingual Programme	Indonesia
Puspa Rani	University of Texas	Indonesia
Heather Tomlinson	N/A	Indonesia/USA
Christine Smith Ellison	University of Ulster	Ireland
Takako Yuki	JICA Research Institute	Japan
Yuriko Kameyama	JICA Research Institute	Japan
Aisha Alsheikh	Injaz Al Arab	Jordan
Abir al Omari	Ministry of Education	Jordan
Mona Jamil Taji	Queen Rania Teacher Academy	Jordan
Khalil Elaian	UNESCO	Jordan
Sammy Khasawneh	University of Jordan, Amman	Jordan
Kabiru Kinyanjui	ADEA	Kenya
John Mugo	Uwezo	Kenya
Barbara Garner Koech	World Vision International	Kenya
Hyekyoung Jung	Korea Institute for Curriculum and Evaluation (KICE)	Korea
Jimin Cho	Korea Institute for Curriculum and Evaluation (KICE)	Korea
Jin Huh	Korea Institute for Curriculum and Evaluation (KICE)	Korea
Minhee Seo	Korea Institute for Curriculum and Evaluation (KICE)	Korea
Sujin Kim	Korea Institute for Curriculum and Evaluation (KICE)	Korea
Bangran Ryu	Korean Educational Development Institute (KEDI)	Korea
Baul Chung	Korean Educational Development Institute (KEDI)	Korea
Byung Young Park	Korean Educational Development Institute (KEDI)	Korea
Chol Kyun Shin	Korean Educational Development Institute (KEDI)	Korea
Heung Joo Kim	Korean Educational Development Institute (KEDI)	Korea
Hu-nam Lim	Korean Educational Development Institute (KEDI)	Korea
Hyeseung Chang	Korean Educational Development Institute (KEDI)	Korea
Hye-sook Shin	Korean Educational Development Institute (KEDI)	Korea
Hye-Young Park	Korean Educational Development Institute (KEDI)	Korea
Hyun-jung Im	Korean Educational Development Institute (KEDI)	Korea
Jae-Boon Lee	Korean Educational Development Institute (KEDI)	Korea
Jee-young Nangung	Korean Educational Development Institute (KEDI)	Korea
Jeongwon Kim	Korean Educational Development Institute (KEDI)	Korea
Ji-eun Kim	Korean Educational Development Institute (KEDI)	Korea

Name	Organization	Location
Jin Hee Kim	Korean Educational Development Institute (KEDI)	Korea
Ju-Ah Kim	Korean Educational Development Institute (KEDI)	Korea
Kangjoo Lee	Korean Educational Development Institute (KEDI)	Korea
Kyungeun Chae	Korean Educational Development Institute (KEDI)	Korea
Meesook Kim	Korean Educational Development Institute (KEDI)	Korea
Mikyung Chung	Korean Educational Development Institute (KEDI)	Korea
Sang-Duk Choi	Korean Educational Development Institute (KEDI)	Korea
Seung-Sil Yang	Korean Educational Development Institute (KEDI)	Korea
Sung Bin Moon	Korean Educational Development Institute (KEDI)	Korea
Sung-Chang Roo	Korean Educational Development Institute (KEDI)	Korea
Sun-Geun Baek	Korean Educational Development Institute (KEDI)	Korea
Sung-Ho Park	Korean Educational Development Institute (KEDI)	Korea
Yang-boon Kim	Korean Educational Development Institute (KEDI)	Korea
Young-Hae Kang	Korean Educational Development Institute (KEDI)	Korea
Jiyoung Kim	Myoungil Middle School	Korea
Giyoun Kim	Seongsu High School	Korea
Larisa Marchenko	Department for Monitoring and Strategic Planning, Ministry of Education, Kyrgyz Republic	Kyrgyz Republic
Abdelsalam Ibrahim Abdallah	Ministry of Education	Libya
Zaid Ben Mohammad Mahmoud	Ministry of Education	Mauritania
Ahmad Aouzi	Mohammad V University	Morocco
Sara Danc	Save the Children	Nepal
Benjamin A. Ogwo	State University of New York, Oswego	Nigeria
Amaka Ezeanwu	University of Nigeria, Nsukka	Nigeria
Benadeth N. Ezekoye	University of Nigeria, Nsukka	Nigeria
Cajethan U. Ugwuoke	University of Nigeria, Nsukka	Nigeria
Chiamka Chukwuone	University of Nigeria, Nsukka	Nigeria
Felix, C. Nwaru	University of Nigeria, Nsukka	Nigeria
Ifeanyi B. Ohanu	University of Nigeria, Nsukka	Nigeria
Ogbonnaya O.Eze	University of Nigeria, Nsukka	Nigeria
Olawale O.Olaitan	University of Nigeria, Nsukka	Nigeria
Mohammad al Yaqoubi	Ministry of Education	Oman
Talat Mahmud	Save the Children	Pakistan
Soofia Aziz	Plan International	Pakistan
Aliya Abbasi	The Aga Khan University Examination Board (AKU-EB)	Pakistan
Isbah Mustafa	The Aga Khan University Examination Board (AKU-EB)	Pakistan
Rana Jilani	The Aga Khan University Examination Board (AKU-EB)	Pakistan
Rooshi Abidi	The Aga Khan University Examination Board (AKU-EB)	Pakistan
Sayem Hanif	The Aga Khan University Examination Board (AKU-EB)	Pakistan
Zehra Abidi	The Aga Khan University Examination Board (AKU-EB)	Pakistan
Juan León	GRADE	Perú
José Rodríguez	Pontificia Universidad Católica del Perú	Perú
Juan Fernando Vega	Pontificia Universidad Católica del Perú	Perú
Cesar Saldarriaga	Plan International	Perú
Esker Copeland	Qatar Foundation	Qatar

Name	Organization	Location
Vladimir Gorshkov	Center for education quality assessment and control of Yaroslavl region, Russia	Russia
Marina Kuznetsova	Center for general education quality assessment at the Russian Academy of Education, Russia	Russia
Anastasia Maximova	Center for International Cooperation In education Development	Russia
Daria Ovcharova	Center for International Cooperation In education Development	Russia
Olga Loginova	Education center at the Prosvyascheniye publishers, Russia	Russia
Andrey Ivanov	Eurasia Heritage Foundation, Russia	Russia
Klara Krasnyanskaya	Laboratory of Mathematics Didactics and Center for General Education Quality Assessment at the Russian Academy of Education, Russia	Russia
Marina Tryakshkina	Social ratings center at RIA news	Russia
Nataliya Tyurina	Social ratings center at RIA news	Russia
Oxana Rydze	The Institute of content and teaching methods of the Russian Academy of Education, Russia	Russia
Ruth Mbabazi	VSO Rwanda	Rwanda
Lea Salmon	LARTES	Senegal
Junko Miyahara	ARNEC	Singapore
S.M. Naicker	Curriculum Development	South Africa
Ceferino Artiles	Universidad de las Palmas de Gran Canaria (ULPGC)	Spain
Celia Morales	University of La Laguna, Faculty of Psychology	Spain
Christian Peake	University of La Laguna, Faculty of Psychology	Spain
Desirée González	University of La Laguna, Faculty of Psychology	Spain
Eduardo García Miranda	University of La Laguna, Faculty of Psychology	Spain
Elaine Bisschop	University of La Laguna, Faculty of Psychology	Spain
Isaac Marco	University of La Laguna, Faculty of Psychology	Spain
Isabel O'Shanahan	University of La Laguna, Faculty of Psychology	Spain
Natalia Suárez Rubio	University of La Laguna, Faculty of Psychology	Spain
Rebeca Villarroel	University of La Laguna, Faculty of Psychology	Spain
Remedios Guzmán	University of La Laguna, Faculty of Psychology	Spain
Verónica Rodríguez	University of La Laguna, Faculty of Psychology	Spain
Yesica Jiménez	University of La Laguna, Faculty of Psychology	Spain
Fayza Khalf Allah	Ministry of Education	Sudan
Jean Bourdon	IREDU CNRS	Switzerland
Renato Opertti	UNESCO International Bureau of Education	Switzerland
Usmon Boboev	Ministry of Education, Republic of Tajikistan	Tajikistan
Hamidou Boukary	ADEA	Tanzania
Bhinyo Panijpan	Institute for Innovative Learning, Mahidol University	Thailand
Patcharin Panjaburee	Institute for Innovative Learning, Mahidol University	Thailand
Wannapong Triampo	Institute for Innovative Learning, Mahidol University	Thailand
Pisarn Soydhurum	Institute for the Promotion of Teaching Science and Technology (IPST)	Thailand
Peter J. Foley	Nakomratchasima College	Thailand
Dr Handoko	SEAMEO Secretariat	Thailand

Name	Organization	Location
Dusita Hrawanchid	SEAMEO Secretariat	Thailand
Piyapa Su-angavatin	SEAMEO Secretariat	Thailand
Shazril Helmi Samsudin	SEAMEO Secretariat	Thailand
Antony Tam	UNESCO Bangkok	Thailand
Aurelie Acoca	UNESCO Bangkok	Thailand
Gwang-Chol Chang	UNESCO Bangkok	Thailand
Huong Le Thu	UNESCO Bangkok	Thailand
Jennie Ekedahl	UNESCO Bangkok	Thailand
Lay Cheng Tan	UNESCO Bangkok	Thailand
Mami Umayahara	UNESCO Bangkok	Thailand
Marlene Cruz	UNESCO Bangkok	Thailand
Min Bahadur Bista	UNESCO Bangkok	Thailand
Ramya Vivekanandan	UNESCO Bangkok	Thailand
Satoko Yano	UNESCO Bangkok	Thailand
Stella Yu	UNESCO Bangkok	Thailand
Ushio Miura	UNESCO Bangkok	Thailand
Chemba Raghavan	UNICEF	Thailand
Tanaporn Perapate	UNICEF	Thailand
Cliff Meyers	UNICEF	Thailand
Monia Raies Mghirbi	Arab League Educational, Cultural and Scientific Organization (ALECSO)	Tunisia
Nejib Ayed	Arab League Educational, Cultural and Scientific Organization (ALECSO)	Tunisia
Salim Kacem	Arab League Educational, Cultural and Scientific Organization (ALECSO)	Tunisia
Chadia Mhirsi	CNIPRE	Tunisia
Faten al Maddah	CNIPRE	Tunisia
Souad Abdelwahed	CNIPRE	Tunisia
Khamees Ben Hamida	Independent Consultant	Tunisia
Noureddin Alsassy	Independent Consultant	Tunisia
Omran Boukhari	Independent Consultant	Tunisia
Najoua Ghriss	Ministry of Education	Tunisia
Mohammad Kamal Essid	Ministry of Education	Tunisia
Atef Gadhoumi	National Center for Technology in Education	Tunisia
Ana Nieto	Dubai Cares	UAE
Beau Crowder	Dubai Cares	UAE
Abdallah al Lughani	Ministry of Education	UAE
Mumbogwe Denis	Annual Inter-Institution Convention (AIIC)	Uganda
Waako Mwite Patrick	Light for All High school	Uganda
Fagil Mandy	Uganda National Education Board (UNEB)	Uganda
Amir Jones	CfBT	United Kingdom
Jacqui Mattingly	CfBT	United Kingdom
Susy Ndaruhutse	CfBT	United Kingdom
Anna Hakobyan	CIFF	United Kingdom
Euan Wilmshurst	CIFF	United Kingdom
Faith Rose	CIFF	United Kingdom
Lucy Heady	CIFF	United Kingdom
Peter Colenso	CIFF	United Kingdom
Sally Faiz	CIFF	United Kingdom
Shikha Goyal	CIFF	United Kingdom
Ed Barnett	DFID	United Kingdom

Name	Organization	Location
Laura-Ashley Boden	DFID	United Kingdom
Carew Treffgame	Institute of Education, University of London	United Kingdom
Charlotte Nussey	Institute of Education, University of London	United Kingdom
Chris Yates	Institute of Education, University of London	United Kingdom
Elaine Unterhalter	Institute of Education, University of London	United Kingdom
Gemma Moss	Institute of Education, University of London	United Kingdom
Helen Longlands	Institute of Education, University of London	United Kingdom
Rosie Peppin Vaughan	Institute of Education, University of London	United Kingdom
Roy Carr Hill	Institute of Education, University of London	United Kingdom
Anita Reilly	Plan International	United Kingdom
Leila Reid	The Great Initiative	United Kingdom
Camilla Addey	University of East Anglia, UK	United Kingdom
Kenneth King	University of Edinburgh	United Kingdom
Aubryn Ally Sidle	Advancing Girls' Education (AGE) in Africa	USA
Robin Vitucci	AFT	USA
Li-Ann Kuan	AIR	USA
Kimberly King-Jupiter	Albany State University/ADEA	USA
Jane T. Benbow	American Institutes for Research	USA
Matthew Williams	AusAID	USA
Milkah Kihunah	Care	USA
Eugene Da	CARE USA	USA
Ginny Kintz	CARE USA	USA
Joyce Adolwa	CARE USA	USA
Lotte Renault	CARE USA	USA
Milkah Kihunah	CARE USA	USA
Mphatso Mlia	CARE USA	USA
Tom Haslett	Central Square Foundation	USA
Steven J. Klees	College of Education, University of Maryland (UMD)	USA
Cassie Landers	Columbia University	USA
Amrita Ahuja	Douglas B. Marshall Jr. Family Foundation	USA
Jeff Davis	EDA	USA
Emily Morris	EDC	USA
April Hattori	EdisonLearning	USA
Heath Branham	EdisonLearning	USA
Jodie Beckley	EdisonLearning	USA
Michael McEachran	EdisonLearning	USA
Nicole Wood	EdisonLearning	USA
Sabrena Davis	EdisonLearning	USA
Tim Ulmer	EdisonLearning	USA
Alexandra Buerkle	ELMA Philanthropies	USA
Tressa Johnson	ELMA Philanthropies	USA
Ben Sylla	FHI 360	USA
Kristen Molyneaux	Firelight Foundation	USA
Dana Schmidt	Hewlett Foundation	USA
Alexandra Draxler	Independent Consultant	USA
Marlaine Lockheed	Independent Consultant	USA
Andrea Rugh	Independent Research Professional	USA
Lisa Deters	Independent Researcher	USA
Linda Lehmil	Independent Researcher/Consultant	USA
Arianna Sloat	INEE	USA

Name	Organization	Location
Kerstin Tebbe	INEE	USA
Lori Heninger	INEE	USA
Maria Lucia Uribe	INEE	USA
Noemi Gerber	INEE	USA
Richard C. Seder	McREL	USA
Sena Pierce	McREL (Hawaii)	USA
Amita Chudgar	Michigan State University	USA
Greg Butler	Microsoft	USA
Frank Method (retired)	N/A	USA
Michaela Reich	OAS	USA
Sweta Shah	Plan International	USA
Cynthia Lloyd	Population Council	USA
Amber Gove	Research Triangle Institute	USA
Emily Miksic	Research Triangle Institute	USA
Tony Baker	RESULTS Education Fund	USA
Molly Eberhardt	Results for Development Institute	USA
Heather Simpson	Save the Children	USA
Pablo Stansbery	Save the Children	USA
Lindsey Perry	Southern Methodist University	USA
Sarah Rizk	Teachers College	USA
Anna Lucia D'Emilio	UNICEF	USA
Robert Fuderich	UNICEF	USA
Susan Durston	UNICEF	USA
Aaron Benavot	University at Albany - SUNY	USA
Dan Wagner	University of Pennsylvania Graduate School of Education	USA
Mitch Kirby	USAID	USA
Natasha de Marcken	USAID	USA
Pape Nomar Sow	USAID	USA
Patrick Collins	USAID	USA
Patrick G. Collins	USAID, Burea for Economic Growth, Education and Environment (E3)	USA
Toon Fuderich	USAID, Office of Afghanistan and Pakistan Affairs (OAPA)	USA
Joyce Malombe	Wellspring Advisors	USA
Dana Schmidt	William and Flora Hewlett Foundation	USA
Kate Torre	Winrock	USA
Meredith McCormac	World Learning	USA
Sarah Harekost	World Learning	USA
Jennifer Daniel	World Vision International	USA
Micael Olsson	World Vision International	USA
Barbara Chilangwa	Camfed International	Zambia
Erison Huruba	Camfed Zimbabwe	Zambia
Edith Ngoma	CHILDFUND - Zambia	Zambia
Francis Katungulu	David Kaunda High School	Zambia
Miyanda Kwambwa	Embassy of Ireland	Zambia
Chikuba Beenzu F.	Examinations Council of Zambia	Zambia
Meldah M. Malungo	Examinations Council of Zambia	Zambia
Lumba Siyanga	FAWEZA	Zambia
Mwazanji P. Thorncroft	FAWEZA	Zambia
Chingumbe Kabanda Richards	Ministry of Education	Zambia

Name	Organization	Location
Sendoi Mulonda	Ministry of Education - Mongu	Zambia
Stanley Mwambazi	Ministry of Education - Nakonde	Zambia
Naomi Mwanza	Ministry of Education - Samfya	Zambia
Lastone Silungwe	Ministry of Education - Kasama	Zambia
Alice Kamwi	Ministry of Education - Mbala	Zambia
Nicodemus Kabungo	Ministry of Education - Mbala	Zambia
Sandala Makumbu	Ministry of Education - Mongu	Zambia
Beatrice M. Botha	Ministry of Education - Samfya	Zambia
Lawrence Musonda	N/A	Zambia
Francis Sampa	Read to Succeed	Zambia
William Kapambwe	Read To Succeed	Zambia
Victor Kabwe	Room to Read Zambia	Zambia
Thomas Alvarez	STS	Zambia
Paul Machona	Time to Learn	Zambia
Chilumba Musonda Nalwamba	UNICEF	Zambia
Stephen Chanda	UNICEF	Zambia
Cornelius Chipoma	USAID	Zambia
Mei Mei Peng	USAID	Zambia
Agnes Mando Banda	Zambian Open University	Zambia
Grace B. Manyonga	ZANEC	Zambia

Existing Measures Subgroup	
Abbie Raikes	UNESCO, France
Alejandro Gomez Palma	Organisation for Economic Co-operation and Development, France
Antoine Marivin	PASEC - CONFEMEN, Senegal
Camilla Addey	British Council, France
Carol Armistead Grigsby	Grigsby Global Strategies, United States
Chedia Belaïd Mhirsi	Centre National d'Innovation Pedagogique et de Researches en Education, Tunisia
Ina V.S. Mullis	TIMSS and PIRLS International Study Center, Boston College, United States
Jean-Marc Bernard	GPE Secretariat, United States
Jennifer DeBoer	Massachusetts Institute of Technology, USA
Jimin Cho	Korean Institute for Curriculum Evaluation, South Korea
Kateryna Shalayeva	International Development and European Law Consultant, Ukraine/ France
Keith Lewin	University of Sussex, United Kingdom
Magdalena Janus	McMaster University, Canada
Markus Broer	American Institutes for Research, United States
Michael O. Martin	TIMSS and PIRLS International Study Center, Boston College, United States
Michael Ward	Organisation for Economic Co-operation and Development, France
Moritz Bilagher	UNESCO, Santiago
Pierre Petignat	Université HEP-BEJUNE, Switzerland
Santiago Cueto	GRADE, Peru
Sangwook Park	Korean Institute for Curriculum Evaluation, South Korea
Savitri Bobde	ASER Center, India
Silvia Montoya	Buenos Aires City Government, Argentina

Consultation in Bangkok, Thailand
Hosted by UNESCO Asia-Pacific Regional Bureau for Education (UNESCO Bangkok), Southeast Asian Ministers of Education Organization (SEAMEO), UNICEF Regional Office for East Asia and the Pacific (UNICEF EAPRO)

Consultation in Myrtle Beach, South Carolina
Hosted by CARE USA

Consultation in Nsukka, Nigeria
Hosted by the University of Nigeria Department of Vocational Teacher Presentation

Consultation in Tunis, Tunisia
Hosted by ALECSO

Consultation in Toronto, Canada
Hosted by OISE/University of Toronto

Consultation in the Canary Islands, Spain
Hosted by University of La Laguna

Working Group meeting in Montreal, Canada
Hosted by UNESCO Institute for Statistics

Annex B: Methodology

The Measures and Methods working group builds on the work of the Standards Working Group convened in May–December 2012. In Phase I of the LMTF, this group proposed a broad definition of learning encompassing seven domains. The task force deliberated on the working group's recommendations at an in-person meeting in September 2012 and decided upon the following seven domains in which the subsequent working group would investigate measurement: physical well-being, social and emotional, culture and the arts, literacy and communication, learning approaches and cognition, numeracy and mathematics, and science and technology. (See *Toward Universal Learning: What Every Child Should Learn*).

The second working group addressed the question of how learning in these domains could be measured. The Measures and Methods Working Group was charged with developing a framework for measuring learning that is both informative to global policy dialogues and relevant to national education goals, preventing a "one-size-fits-all" approach. The group was instructed that they could consider measures that correspond to aspirational goals and not focus too much on the resources needed to implement the assessments, as this would be covered by the third and final working group on implementation. However, the working group did keep feasibility in mind when designing these approaches.

In addition to the overarching question on how learning could be measured, several other questions were posed to the working group by the task force after its first meeting in September 2012 regarding the scope of the LMTF recommendations:

- **Should learning be measured for all children and youth or only those who are in school?** Universal primary education has been achieved or is close to being achieved in the majority of countries (UNESCO 2012). As schools are the primary vehicles for improving learning outcomes, some argue that learning assessments should be conducted only within schools to simplify and focus on making improvements to the system. Others cited low enrollment numbers in pre-primary programs (48% gross enrollment ratio, GER, globally) and secondary school (70% GER globally) (UNESCO 2012), especially in low- and middle-income countries as reasons why the recommendations should extend to children outside of formal school settings. A related question is whether learning should be measured in nonformal learning programs or only in formal schooling systems.

- **Should learning be measured by age cohort or grade level?** The discussion on age cohort and/or grade level is related to the discussion on whether the recommendations should apply to all children or only children in school. Some argue that an age-based model would keep governments accountable for learning of all children, whether or not they are enrolled in school. Others argue that the varying ages at which children begin school globally would make grade levels a more pragmatic way of measuring learning, especially in any internationally comparable way and also would allow connecting the learning measures to attributes of the school trajectories that can be addressed by policies and interventions.

The Second Working Group on Measures and Methods

The second phase of the LMTF began with convening the Measures and Methods Working Group. The LMTF Secretariat issued an open call for applications in August 2012 and sent invitations to targeted individuals and other entities with substantiate experience in learning measuring, including all the regional assessment outfits. The Measures and Methods Working Group consists of 57 members in 27 countries (Afghanistan, Argentina, Australia, Brazil, Canada, Colombia, Ecuador, France, Ghana, India, Kenya, Korea, Pakistan, Papua New Guinea, Peru, Qatar, Russia, Rwanda, Senegal, South Africa, Spain, Switzerland, the Netherlands, Turkey, Ukraine, United Kingdom, United States), with equal representation from the Global North and Global South.[27] César Guadalupe (Ed.D.), a lecturer and researcher at Universidad del Pacífico in Lima, and a nonresident fellow of the Brookings Institution, serves as chair of the working group.

Before that, Guadalupe led the Learning Outcomes Section at the UNESCO Institute for Statistics. He also served as UIS' regional advisor for Latin America and the Caribbean. He has held the positions of head of education statistics and head of the Analysis Unit in the Peruvian Ministry of Education and has extensive experience in carrying out social research to improve the responsiveness to policy needs, and developing policy-oriented analytical and research activities within subnational, at national and international levels.

Guadalupe and the Secretariat facilitated the Measures and Methods Working Group from October 2012 to March 2013. Similarly to Phase I of the project, the working group was divided into three subgroups— Early Childhood, Primary, and Lower Secondary. The subgroups worked virtually by completing assignments and participating in teleconferences, email discussions, and small group discussions.

The working group completed a preliminary mapping of current assessments in each of the seven domains (including measures at the international, regional, national, and sub-national levels) and developed a document "Prototype Framework for Measuring Learning Outcomes" with proposed goals and indicators for measuring learning based on the current availability of measures (see annex D).
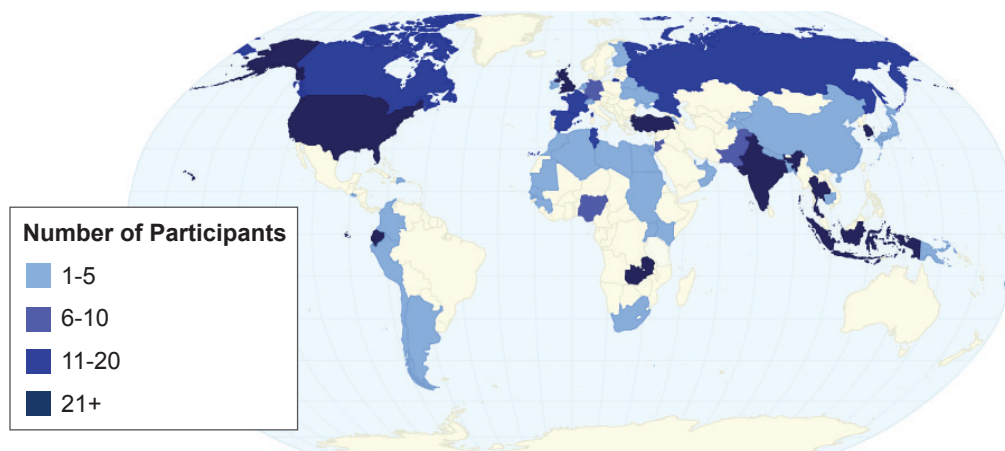
Second Public Consultation Period

The working group's "Prototype Framework for Measuring Learning Outcomes" was disseminated for public consultation between December 2012 and January 2013. This preliminary document was circulated along with a consultation toolkit and a brief overview and status update of the LMTF project, to help frame and guide the discussions around how learning should be measured in the seven domains identified in Phase I. Consultation documents were made available in Arabic, English, French, Russian and Spanish. More than 600 people in at least 57 countries provided feedback by either participating in an in-person consultation or submitting comments electronically. Figure 5 shows the geographic representation of participants in the measures and methods consultation period. Table 4 lists the countries and approximate number of participants.

Several overarching themes emerged from the Phase II consultations:

- Consultees were supportive of a hybrid model, where a small set of domains would be measured globally (i.e. in every country) and other domains would be measured at the national level as determined by country priorities. For example, while many suggested that a global measurement for reading within the Literacy and Communication domain is feasible and desirable, a measurement for creativity within the Culture and the Arts domain would be feasible and more relevant if measured nationally or even locally.

- The prototype framework proposed two levels of learning at the global level—minimum and desirable. While the idea of multiple levels of learning was welcomed, measurement on a continuum of levels, or progress across a benchmarking scale was advised, rather than only at minimum and desirable levels. The feedback from the consultation indicated that a benchmarked scaled would be appropriate to prevent setting lower standards; where the "minimum" would become the acceptable standard. The feedback from the public consultation was supportive of a system of progressively more difficult benchmarks, even benchmarks below the levels currently measured in internationally comparable assessments. However, there was no support for setting international benchmarks for only high- and upper-middle-income countries, as in currently the case, since low- and lower-middle-income countries are not participating in internationally benchmarked assessments for various reasons.

- The feedback from the public consultation supported the demand for measurement of 21st-century skills and competencies related to global citizenship; however, there was no consensus on the definition of these competencies. There was general concern that 21st-century skills or global citizenship would be very difficult to measure in a standardized way, and even more so, on a global scale. As a starting point, it was recommended that the working group starts by defining the components of 21st-century skills and global citizenship, in order to avoid assumptions of greater breadth than actually intended.

- There were many different reactions around the progressive development between Phase I and Phase II of the LMTF project. Many felt that the Measures and Methods Working Group needed to better align recommendations for measurement with the seven domains of learning from Phase I. There was also a consensus that while the measures in the Literacy and Communication and Numeracy and Mathematics domains are important and should continue to be used, there is also a need to move beyond measuring just these two domains.

- There was a request for alignment of terminology, specifically around the goals, measures and indicators, and benchmarks, to ensure that they are used consistently through the prototype framework and the remaining LMTF documentation. There was also agreement that the differences amongst these terms should be analytically useful, reasonably transparent for non-technical policymakers, and yet scientifically sound. It was also suggested that a glossary of words be developed and used consistently across language translations.

Figure 5. Countries Represented in the Measures and Methods Consultation Period

Number of Participants
- 1-5
- 6-10
- 11-20
- 21+

Table 4. Countries and Participants Represented in the Measures and Methods Consultation Period*

Region	Participating Countries	Region	Participating Countries
Eastern Africa	• Kenya (2) • Rwanda (1) • Sudan (1) • Tanzania (11) • Uganda (3) • Zambia (30)	Southeastern Asia	• Cambodia (1) • Indonesia (37) • Papua New Guinea (1) • Singapore (1) • Thailand (25)
Northern Africa	• Algeria (1) • Egypt (2) • Libya (1) • Mauritania (1) • Morocco (1) • Tunisia (13)	Southern Asia	• Bangladesh (3) • India (26) • Nepal (1) • Pakistan (8)
Western Africa	• Guinea (1) • Nigeria (9) • Senegal (3)	Eastern Europe	• Armenia (2) • Belarus (1) • Kyrgyz Republic (1) • Russian Federation (12) • Tajikistan (1) • Turkey (25) • Ukraine (1)
Southern Africa	• South Africa (1)		
Central America	• El Salvador (1)		
North America	• Canada (14) • United States (148)	Northern Europe	• Finland (1) • Ireland (1) • United Kingdom (32)
South America	• Argentina (1) • Chile (1) • Colombia (4) • Ecuador (23) • Peru (2)	Western Europe	• France (13) • Germany (10) • Netherlands (1) • Spain (14) • Switzerland (4)
Caribbean	• Dominican Republic (1)	Middle East	• Jordan (7) • Oman (1) • Qatar (1) • United Arab Emirates (3)
Eastern Asia	• China (1) • Korea (38) • Japan (2)	Oceania	• Fiji (1)

Measures and Methods Working Group Meeting

Twenty members of the Measures and Methods Working Group and LMTF Secretariat convened on January 22–23, 2013, in Montreal to consider the reactions and feedback from the public consultation, to address the questions regarding the scope of the project, and to develop one or more approaches for measuring learning that would be presented to the task force on February 20–21, 2013.

The working group defined early childhood as the time from birth through the age of compulsory primary school entry. At the global level, they agreed that learning should be measured ideally in households, as the majority of children worldwide are not enrolled in pre-primary programs; but, in cases where almost all children start school on time, it would be pragmatic to assess school readiness in the early grades of primary.

For primary and lower secondary, the working group recommended using the 1997 revision of the International Standard Classification of Education (ISCED), thus ISCED97 Level 1 for primary and ISCED97 Level 2 for lower secondary. While the working group reached consensus that at the global level measurement should be in schools, they also agreed that they would provide guidance for measurement in contexts in which lower secondary is not compulsory, where primary/lower secondary enrollment levels are low and in cases where children and youth are displaced by conflict or disaster. The working group also agreed to provide guidance for how countries can decide on measuring learning in nonformal programs.

The working group also considered two approaches to measurement of learning globally to present to the task force, from which the task force could select elements that are feasible and desirable for measurement at the global scale. The approaches for measurement discussed were (1) expanding existing international comparable assessments to a global scale; (2) expanding national capacity and improving decisionmaking for measuring learning.

The *International Approach* describes how current internationally comparable assessments could be scaled up to a global level. The advantage of this approach is that learning would be measured on a common scale that is comparable across countries allowing for measuring global progress using common benchmarks, so that, for example, "proficiency in reading" would be not only defined in a common way, but measure using the same metric. The limitation of this approach is that the available measures are concentrated primarily around primary and lower secondary reading, mathematics and science, with some promising practices in early childhood (across five domains), information and communication technology, and citizenship. Learning outcomes for out-of-school children and youth would not be measured in this model, even though, the concurrent use of access and completion measures can preempt potential bias in interpreting the information. The assessments and benchmarks were developed on a scale suitable for those countries currently implementing these measures (mostly OECD or middle-income countries) which might lead to a "floor effect" whereby a large proportion of students would appear showing very low levels of proficiency undistinguishable from random guessing. This risk would be minimized if the characteristics of the countries that are new to the measurement endeavors are taken into account in their regular redesign and adjustment of the frameworks and instruments.

Finally, a very specific set of common measures would facilitate the process of placing the educational challenges in the forefront of the international development agenda.

The *National Approach* describes how the international education community could support national capacity and decisionmaking. The advantage of this approach is that countries unwilling or unable to join the internationally comparable assessments could still measure progress toward global learning goals. Countries could use measures that capture the range of abilities in their countries and include out-of-school children and youth if desired. The disadvantage of this approach is that the data would not be comparable at the global level and therefore measuring progress toward a global goal could not be done in a robust manner since "similar" benchmarks would mean something different across countries. A construct such as "proficiency in mathematics" would be understood and measured differently from country to country and therefore there would be no way to judge which country or sub-population is getting closer to the global goals.

The Secretariat and Measures and Methods Working Group identified 10 indicators that respond to the learning standards that are feasible and/or have broad support from the consultation that could be proposed for a set of global learning goals. The task force will decide which indicators are politically practical and actionable. See below for a description of these op-

tions. The 10 indicators that are both feasible and/or are supported by the consultation are in relation to the following:

- *Early childhood:* (1) school readiness (across five domains: physical well-being, social and emotional, literacy and communication, numeracy and mathematics, learning approaches and cognition)

- *Primary:* (2) social and emotional competencies, (3) reading, (4) mathematics, (5) science

- *Lower secondary:* (6) citizenship, (7) reading, (8) collaborative problem solving, (9) mathematics, (10) science

In addition to the learning metrics proposed below, the working group recommended that any global learning metrics be paired with access metrics so that the millions of children and youth who are not in school are not excluded from this framework.

The working group and Secretariat identified what is technically feasible and desirable for global learning metrics. They then asked task force to decide which of the metrics on this list, or any other metrics, are politically practical and actionable to propose for the post-2015 development goals. The results are given in table 5.

Table 1. Feasible and Desirable Learning Metrics at the Global Level (Proposed to Task Force at 20-21 February LMTF Meeting in Dubai)

Stage	Access indicator	Learning indicator	Learning Metrics for Approach 1	Learning Metrics for Approach 2	Technically feasible to scale up now?	Broad support from consultation?
Early childhood	On-time entrance to primary (net intake rate)	School readiness across five domains (physical well-being, social and emotional, literacy and communication, learning approaches and cognition, numeracy and mathematics)	EDI: % of children in the "ready" or "very ready" percentiles (25th–100th percentiles); % reduction in disparities between groups (e.g. wealth quintiles, gender, urban/rural)	EDI or other technically sound school readiness instrument: % of children ready for school; % reduction in disparities between groups (e.g., wealth quintiles, gender, urban/rural)	✓	✓
Primary	On-time completion of primary;[28] Population age 15–19[29] that has completed primary school	Social and emotional	None identified.	Strengths and Difficulties Questionnaire (SDQ) or other technically sound assessment; % reduction in disparities between groups on measure of lower secondary literacy; % increase in proficiency over time		✓
		Reading	PIRLS or pre-PIRLS: % of students reading at PIRLS level 4; % reduction of disparities between groups; % meeting intermediate benchmarks; % increase in proficiency over time	PIRLS or pre-PIRLS, regional assessments, or technically sound national assessments: % reduction in disparities between groups on measure of reading; % increase in proficiency over time	✓	✓
		Mathematics	TIMSS or TIMSS-numeracy: % of students proficient in TIMSS level 4; % reduction of disparities between groups; % meeting intermediate benchmarks; % increase in proficiency over time	TIMSS, TIMSS-numeracy, regional assessments, or technically sound national assessments: % reduction in disparities between groups on measure of primary mathematics; % increase in proficiency over time	✓	✓

Table 1. Feasible and Desirable Learning Metrics at the Global Level (Proposed to Task Force at 20-21 February LMTF Meeting in Dubai)

Stage	Access indicator	Learning indicator	Learning Metrics for Approach 1	Learning Metrics for Approach 2	Technically feasible to scale up now?	Broad support from consultation?
		Science	TIMSS: % of students proficient in TIMSS level 4; % reduction of disparities between groups; % meeting intermediate benchmarks; % increase in proficiency over time	TIMSS, regional assessments, or technically sound national assessments: % reduction in disparities between groups on measure of primary science; % increase in proficiency over time	✓	
Lower secondary	Transition rate to lower secondary	Citizenship	ICCS:	ICCS or national assessment of social or citizenship knowledge and values: % reduction in disparities between groups; % increase in proficiency over time	✓	✓
	On-time completion of lower secondary	Literacy	PISA:	PISA or technically sound national assessment: % reduction in disparities between groups on measure of lower secondary literacy; % increase in proficiency over time	✓	✓
	Population age 20-24[30] that has completed primary school	Collaborative problem solving/ Critical thinking	None identified	ATCS21S Collaborative Problem Solving Assessment or other measure: % reduction in disparities between groups on measure of lower secondary literacy; % increase in proficiency over time		✓
		Mathematics	TIMSS or PISA:	TIMSS, PISA or technically sound national assessment: % reduction in disparities between groups on measure of lower secondary mathematics; % increase in proficiency over time	✓	✓
		Science	TIMSS or PISA:	TIMSS, PISA or technically sound national assessment: % reduction in disparities between groups on measure of lower secondary science; % increase in proficiency over time	✓	

Annex C: Participants in Learning Metrics Task Force Meeting on Measures and Methods

Learning Metrics Task Force
Meeting on Measures and Methods
February 20–21, 2013
Hosted by Dubai Cares at Al Murooj Rotana Hotel, Dubai

Attendees

Task Force Member Organization Representatives:	
Organization	**Representative(s)**
ActionAid International; Global Partnership for Education Board Representative for Northern Civil Society	David Archer, International Head of Education
African Union Commission	Beatrice Njenga, Head of Education Division, Directorate of Human Resources Science and Technology
Arab League Educational, Cultural and Scientific Organization (ALECSO)	Ahmed Sayed Khalil Ahmed, Director of Education Monia Rais Mghirbi, Arab Regional Agenda for Improving Education Quality (ARAIEQ) Project Manager
Association for the Development of Education in Africa (ADEA); Working Group on Implementation Chair	Dzingai Mutumbuka, Chair
Australian Agency for International Development (AusAID)	Raymond Adams, Special Advisor, ACER; Professor, University of Melbourne
Campaign for Female Education (Camfed) International, Representing Southern Civil Society	Lucy Lake, Chief Executive Officer
City of Buenos Aires, Argentina	Mercedes Miguel, Director General of Education Planning
Dubai Cares/United Arab Emirates	H.E. Reem Al-Hashimy, Chair and Minister of State Tariq Al-Gurg, Chief Executive Officer Beau Crowder, Director of Programmes Maria al Qassimi, Country Programme Officer
Education International	David Edwards, Deputy General Secretary and GPE Board Representative for the Teaching Profession
Global Partnership for Education	Jean-Marc Bernard, Senior Education Specialist
Government of Assam, India	Dhir Jhingran, Principal Secretary
International Education Funders Group (IEFG)	Chloe O'Gara, Co-Chair
Korea Institute for Curriculum and Evaluation (KICE)	Seong Taeje, President Ji Min Cho, Head, Office of Global Cooperation Mi Young Song, Research Fellow Min-Hee Seo, Associate Research Fellow
Ministry of Education of Kenya	George Godia, Permanent Secretary

Task Force Member Organization Representatives:	
Organization	**Representative(s)**
Office of the UN Secretary General	Itai Madamombe, Global Education Advisor
Pearson	Michael Barber, Chief Education Advisor (Task Force Co-Chair)
	Amanda Gardiner, Head of International Affairs
Pratham	Rukmini Banerji, Director of Programs (Task Force Co-Chair)
Queen Rania Teacher Academy	Tayseer Al-Noaimi, President; former Jordan Minister of Education
Southeast Asian Minister of Education Organization (SEAMEO)	I. Handoko, Deputy Director for Programme & Development
USAID	John Comings, Education Advisor
UK Department for International Development (DfID)	Ian Attfield, Education Adviser, DfID Tanzania
UNESCO	Maki Hayashikawa, Chief, Section for Basic Education, Division for Basic Learning and Skills Development
	Abbie Raikes, Program Specialist, Early Childhood Education
UNICEF	Jo Bourne, Deputy Executive Director (Programmes)
Working Group on Measures and Methods Chair	César Guadalupe, Associate Researcher, Universidad del Pacífico, Perú
Working Group on Standards Chair	Seamus Hegarty, former chair, International Association for the Evaluation of Educational Achievement (IEA)
World Bank	Marguerite Clarke, Senior Education Specialist

Task Force Secretariat:	
Organization	**Representative(s)**
Center for Universal Education (CUE) at the Brookings Institution	Rebecca Winthrop, Senior Fellow and Director
	Maysa Jalbout, Nonresident Senior Fellow
	Kate Anderson Simons, Technical Lead, Learning Metrics Task Force
	Mari Soliván, Project Manager, Learning Metrics Task Force
UNESCO Institute for Statistics (UIS)	Hendrik van der Pol, Director
	Albert Motivans, Head of Education Indicators and Data Analysis Section
	Maya Prince, Research Assistant

Observers:	
Organization	**Representative(s)**
Australian Council for Education Research (ACER)	Alan Egbert, Manager, ACER Dubai
The Commonwealth Secretariat	Pauline Greaves, Head of Section, Social Transformation Programmes Division
Idara-e-Taleem-o-Aagahi (ITA)	Baela Raza Jamil, Director, Programs
South Asian Association for Regional Cooperation (SAARC)	Prasanna Gamage, Director, Sri Lanka, SAARC Secretariat

Annex D: Prototype Document Released for Consultation Period

Note to readers: The following prototype framework and the indicators, targets and measures proposed in this document are presented to initiate dialogue on how learning could be measured at the global and national levels. This framework is intentionally incomplete. To complete it, the Measures and Methods Working Group seeks the input of stakeholders from around the world through a broad consultation process. Feedback from global consultations will be used to refine the goals, indicators, targets and measures, as the task force works to create a framework that is holistic yet feasible for measurement at the global level. To ensure your feedback informs the working group's final recommendations to the task force, please email your comments to learningmetrics@brookings.edu by January 18, 2013.

What is the Learning Metrics Task Force?

The benefits of education—to individuals and their families, as well as to local, national and international communities in relation to prosperity, health and social stability—are well known, but these benefits are conditional on the levels of learning children actually acquire through schooling. Despite commitments and progress in improving access to education at the global level (i.e., EFA Goals and MDG 2), learning levels are still low. According to the 2012 EFA Global Monitoring Report estimations, at least 250 million primary-school-age children around the world are not able to read, write or count well, including those who have spent at least four years in school.[31] However, even the best possible estimations are not enough to understand the full scale of the crisis, because measurement of learning is limited, and unlike the measurement of access, difficult to assess at the global level.

To advance progress for children and youth around the world, it is critical that learning is recognized as essential for human development. As EFA and the MDGs sunset in 2015, and the UN secretary-general launches Education First, the education sector has a unique window of opportunity to raise the profile of international education goals and ensure that learning becomes a central component of the global development agenda. To do this, the global education community must work collectively to define global ambition on improving learning and practical actions to deliver and measure progress.

In response to this need, UNESCO through its Institute for Statistics (UIS) and the Center for Universal Education (CUE) at the Brookings Institution have co-convened the Learning Metrics Task Force. The overarching objective of the project is to catalyze a shift in the global conversation on education from a focus on access to access plus learning. Based on recommendations of technical working groups and input from broad global consultations, the task force aims to make recommendations to help countries and international organizations measure and improve learning outcomes for children and youth worldwide.

The task force engages high-level political actors, technical experts and practitioners in a year-long global consultation process in order to build consensus around three questions:

- What learning is important for all children and youth?

- How should learning outcomes be measured?

- How can measurement of learning improve education quality?

The task force is comprised of national and provincial governments, regional organizations, multilateral organizations, teachers' organizations, civil society groups and donor agencies.[32]

In Phase I of the project, the Standards Working Group convened from May to October 2012 to make recommendations for what learning is important globally. The prototype recommendations were released for public consultation in August to September 2012 and modified based on feedback from more than 500 individuals in 57 countries. A draft framework was presented to the task force at an in-person meeting in September 2012. Over two days, the task force finalized a framework to be used for the subsequent working group to investigate domains of learning outcomes.

The release of this document for global consultation is part of Phase II of the project. The Measures and Methods Working Group,[33] which began its work in October 2012, has been charged by the task force with developing a framework for measuring learning that is both informative to global policy dialogues and relevant to national education goals. This document presents a "prototype"[34] for how learning could be measured at the global and national levels.

How Does Learning Occur?

Learning occurs constantly in multiple contexts—homes, schools, communities and workplaces—and through various processes, ranging from highly structured, guided instruction to informal, unstructured processes.

The Classification of Learning Activities (CLA) describes different learning opportunities and pathways for education/learning.[35] The CLA framework provides criteria for classifying education and learning into three broad categories: (1) formal education, (2) nonformal education, and (3) informal learning.

In this framework, learning activities must be (1) intentional, where individuals engage in activities with a purpose to improve their knowledge, skills or competence; and (2) organized, so that it involves a transfer of information. Key concepts in the CLA framework include:

- *Intentional learning* is when individuals intentionally/voluntarily search for knowledge, skills, competences or attitudes of lasting value, and that intention is formulated by the learner before starting the activity.

- *Organized learning* is planning learning in a given pattern with "explicit or implicit aims," and involves an agency to set up the learning environment and a method of teaching though which the communication is organized.

- *Formal Education* refers to education provided in systems of schools, colleges, universities and other formal educational institutions that establish the cycle of learning from early childhood through young adulthood (up to the age of 25 years).

- *Nonformal Education* refers to organized and sustained educational activities that can take place within and outside formal educational institutions; however, they typically do not have a defined path (i.e., individuals do not need to complete one level to move on to the next), and hence do not conform to the definition of formal education. These may include but are not limited to literacy programs, life skills, work skills and general culture.

- *Informal Learning* refers to intentional, but less organized and less structured learning activities at home, at work and in everyday life, on a "self-directed, family-directed or socially directed basis."

Wagner, Murphy and De Korne (2012) propose the following framework for describing the spectrum of where and how learning occurs, which also includes unintentional learning:

- *Formal contexts*—highly structured processes: Typically describes learning in primary, secondary and tertiary institutions with trained teachers and formal curricula and materials (e.g., textbooks).

- *Nonformal contexts*—highly structured processes: Refers to contexts not subject to systematic regulations, often including pre-primary programs, religious schools, private tutoring and youth literacy programs. The structure of these programs is similar to instruction in formal contexts.

- *Formal contexts*—unstructured/informal processes: Describes learning that occurs in schools and other formal institutions that is not part of the formal curriculum—for instance, what children and youth learn from their peers.

- *Non-formal contexts*—unstructured/informal processes: Refers to the many learning opportunities that occur in everyday life in the home, community, marketplace, etc.

These two frameworks provide background on the various pathways in which humans acquire knowledge, skills, values and abilities. What learning should be assessed and when it should be assessed is closely related to where and how learning occurs through these various contexts and processes.

Where Are Children and Youth Learning?

Early Childhood: Globally, 164 million children are enrolled in preschool programs, and the pre-primary GER is 48% (UNESCO 2012). However, access to pre-primary programs is highly inequitable, with a GER of only 15% in low-income countries. Children least likely to be enrolled in preschool are those belonging to minority ethnic groups, those with less educated mothers, and those who speak a home language different from the language used in school.[36] These are also the children most likely to benefit from high-quality preprimary programs. While many children, especially in high-income countries, attend formal, regulated preprimary programs, the majority of the world's young children only learn in non-formal contexts through unstructured or informal processes. For these children, learning typically occurs in the home and community through interactions with parents, siblings and other family members. Even when children are enrolled in preprimary programs, they may not be exposed to quality formal early learning opportunities.

Primary: Partially due to a global focus on universal primary education, the majority (89%) of primary age children are now enrolled in school.[37] Free, compulsory primary education is recognized as a fundamental human right,[38] and primary education is compulsory in almost every country.[39] Still, there are nearly 61 million out-of-school children of primary-school age, a number that has stagnated since 2008.[40] While many children are either not enrolled in school or enrolled in second-chance programs, the majority of primary age children globally are learning in formal contexts. However, the degree to which formal processes are good enough to

ensure children's right to a decent education depends in large part on the quality of the teachers, curriculum, and materials found in the school. In schools where there are enough qualified teachers and materials to respond to each individual child's learning needs, academic learning occurs through formal processes. In schools where teachers are not properly qualified, are overextended, or do not come to work regularly, learning still occurs through peer-to-peer interactions, but not necessarily the types of learning intended by the school system.

Postprimary: The category of postprimary refers to the various contexts in which children learn beyond primary schooling. For most children, postprimary refers to secondary education. The UIS reports that in 2010, lower secondary education was part of compulsory education in three out of four countries reporting data, and upper secondary was included in compulsory education in approximately one in four countries.[41] It is estimated that globally, 91% of children who entered school stay there until the end of primary and 95% of those students transition to secondary school.[42]

However, for children in low-income countries, only 59% make it to the last year of primary school and 72% of those students successfully transition to secondary schooling.[43] For children who do not attend secondary school, learning occurs mainly through work, family and community experiences (i.e., non-formal, unstructured contexts).

When Are Children Learning?

The times when children learn can be described through stages (early childhood, primary and postprimary), schooling levels, and/or age groups. How these groupings correspond to one another varies across countries and even across individual children. Table 6 attempts to define the stages, schooling levels and approximate age spans for these groups. The schooling levels are based on the 1997 revision of the International Standard Classification of Education (ISCED).[44] Note that the age spans overlap intentionally to account for wide variations in when children begin and end school.

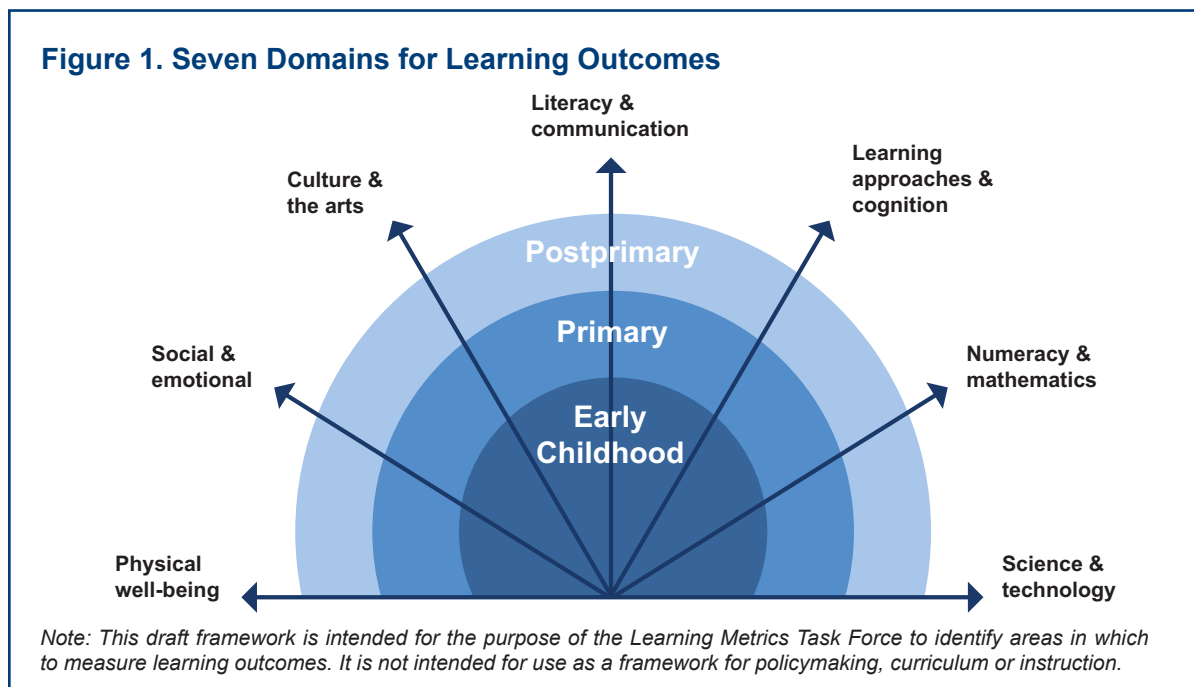| Table 1. Stages, Schooling Levels and Approximate Age Spans for Measuring Learning Outcomes		
Stage	**Schooling Levels**	**Approximate Age Span (years)**
Early childhood	Preprimary, including formal and nonformal (ISCED 0)	0–8
Primary	Lower and upper primary (ISCED 1)	5–15
Postprimary	Lower secondary (ISCED 2)	10–19

What Learning Is Important for All Children and Youth?

Given the various structures, places and times in which humans learn, it is difficult to define what outcomes related to learning are universally important, especially at a global level. However, certain outcomes have been identified as important for all children and youth to develop based on (1) research, (2) global policies and dialogues, and (3) the real-life experience of education practitioners. Based on a research review, recommendations from the Standards Working Group, input from global consultations, and task force deliberations, seven domains and corresponding subdomains of outcomes related to learning were proposed as areas for exploration by the subsequent Measures and Methods Working Group:

- Physical well-being

- Social and emotional

- Culture and the arts

- Literacy and communication

- Learning approaches and cognition

- Numeracy and mathematics

- Science and technology

Figure 6 represents the proposed framework of seven domains. Each arrow in the diagram represents one domain of outcomes, radiating outward as a child expands his or her development or competency in a given area. The segments of concentric circles represent the three stages in which the task force will concentrate its recommendations: early childhood (birth through primary school entry); primary, and postprimary (end of primary through end of lower secondary). The arrows extend outward beyond the diagram to indicate that an individual continues learning more deeply in any given area at the upper secondary, tertiary or technical/vocational level, or through nonformal learning opportunities.

Figure 1. Seven Domains for Learning Outcomes

Note: This draft framework is intended for the purpose of the Learning Metrics Task Force to identify areas in which to measure learning outcomes. It is not intended for use as a framework for policymaking, curriculum or instruction.

Subdomains were proposed by the Standards Working Group as areas in which learning may be measured. For example, the subdomains for early childhood literacy and communication are as follows:

- Stage: Early Childhood

- Domain: Literacy and communication

- Subdomains: Receptive language, expressive language, vocabulary, print awareness

This framework was used as a starting point for conversations about how learning outcomes should be measured. However, working group members and other stakeholders raised the following issues about the framework and the diagram:

- The diagram combines outcomes related to contextual factors (e.g., physical well-being), processes (e.g., learning approaches and cognition), and academic content (e.g., science and technology), and some of the domains include several of these aspects (e.g., within social and emotional, one may learn academic content through the process of teamwork). It was suggested that the working group clarify these distinctions when describing the diagram.

- Humans learn by making connections across content areas and experiences. This diagram may send the message that learning should occur in discrete domains rather than through an integrated curriculum.

- While the diagram rightly places an equal importance on all seven domains, different purposes might lead users to emphasize some more than others at different levels (i.e., global goals might focus on one or two domains, while national goals might focus on more).

- In some domains, subdomains were too extensive, while in other domains they were not comprehensive enough.

- Initially, the intensity of color in the diagram in the early childhood years was meant to represent the intense capacity for brain development that occurs during this time in a child's life.[45] However, several stakeholders pointed out that this implies that learning in one stage is more important than another, and the capacity for brain development does not necessarily link to learning outcomes.

As the task force works toward operationalizing learning in these domains, the seven domains will become less important than the subdomains, and these will be refined taking into account the actual availability of evidence coming from existing measurement endeavors with different attributes (in most cases based on years of research and validation). As the framework for how learning is measured is refined, the working groups will continue to refer back to the original standards framework and develop a rationale for why a particular area of learning is or is not included at subsequent stages in the process.

Proposed Measurement Framework

Taking into account how, where, and when children learn and what learning is important for all children and youth, the Measures and Methods Working Group proposes a framework for assessing learning that has three tiers of learning goals: global minimum, global desirable, and nationally determined. For the global minimum and global desirable goals, every effort would be made to measure learning in these areas in all countries, although not necessarily using internationally comparable assessments in each country.[46] The Learning Metrics Task Force would concentrate its recommendations on learning outcomes in these areas. The nationally determined goals would capture diverse needs, cultures and priorities, and may be used as examples in the recommendations of the task force.

Global minimum goals are:

- Foundational for later learning and development.

- Based in the reality of where the majority of the world is currently in 2012.

- Achieved through public action carried out by a diversity of actors (governments, communal, private) in diverse settings.

- Feasible for measurement at a global level, although existing capacity may need to be improved to measure in all countries.

Global desirable goals are:

- Based on expectations about where the world needs to be in the next 15 to 20 years.

- "Ladders of opportunity"—skills and abilities without which people's opportunities in life are severely limited.

- Typically reached through the formal education system.

Nationally determined goals are:

- Tied to national or local culture.

- Needed more in some countries than others (e.g., HIV/AIDS knowledge; specific types of disaster and safety knowledge).

- Related to dominant characteristics of a country (e.g., multilingual skills).

- Measured only in countries that determine these domains are a priority.

In order to be feasible and measurable at the global level, the global minimum goals would be indicative of learning progress but not comprehensive. The global desirable goals would be more challenging. Using this model, the "global minimum" and "global desirable" goals would be measured at baseline (likely in 2015) in all countries where data are available, and targets would be set for countries to achieve progress toward these goals during the following 15- to 20-year period. In some cases, goals can be measured using a com-

mon metrics to allow for tracking of progress at a global scale. In this model, all countries would ideally measure progress toward both the minimum and the desirable goals, but the targets may differ based on what is feasible for each country. There are multiple options for how goals would be set and measured, for example:

- X% children and youth meet global minimum by 2030 and Y% meet the global desirable goal;

- Reduce by half the number of children not meeting the minimum and desirable goals;

- An X% increase per year in children and youth meeting the global minimum and global desirable goals;

- An equity goal: an X% decrease in the difference between mean scores of advantaged and marginalized groups within a country (e.g., girls, poor children, children in rural areas, children in emergencies).

For countries not currently measuring progress toward these goals:

- The Learning Metrics Task Force could develop a process for ministries, civil society, donors and other stakeholders to determine capacity and demand for measurement within countries.

- Existing measures of learning used at the subnational level could be scaled up or used to get a partial baseline.

- The country could adopt international measures through participation in an existing measurement effort.

- Countries or agencies could use other population metrics (health, educational attainment, etc.) to estimate a baseline.

Finally, it is important to note that while a complete baseline for all countries in 2015 may not be possible, the education community should continue developing better tools for assessment while promoting the development of national capacities to measure and the importance of doing so.

Current Feasibility for Measurement

The Measures and Methods Working Group conducted a preliminary canvass of measures of learning outcomes in the seven domains and many examples are listed here. The following list is by no means exhaustive; the working group hopes that many more measures will be identified through the consultation process. The working group categorized the feasibility for measuring learning outcomes in the following way[47]:

- Some learning outcomes are measured informally or a measure is developed but not standardized.

- Some learning outcomes have a standardized measure available but not used across populations.

- Some learning outcomes have a standardized measure available and used across populations within one or more countries.

- Some learning outcomes are important but developing standardized measures is neither feasible nor desirable.

Based on the work of the Standards Working Group as to which learning outcomes are feasible to measure, supported by research and policy, and desirable by educators, the Measures and Methods Working Group proposes the following framework (table 7) for measuring learning at the global and national levels. The potential targets and measures for these goals are described below.

Table 2. Global and National Goals and Indicators for Learning Outcomes				
Learning Stage	**Global Goal**	**Global Minimum Indicator**	**Global Desirable Indicator**	**Nationally Determined Goals and Indicators**
Early Childhood	**Young children start school on time and are ready to learn**	Children are free from chronic nutritional deficits, developmental difficulties, health conditions, and start school on time	Children demonstrate competence in domains associated with school readiness, including physical well-being, social and emotional, language, numeracy, learning approaches and cognition	Learning in other areas determined priority by national early learning priorities
Primary	**Children acquire basic literacy, numeracy, and global citizenship skills by the end of primary school**	All children complete primary school and acquire basic reading skills by the age of 15	All children complete primary school on time and acquire intermediate skills in literacy, numeracy and mathematics, and global citizenship	Learning in other areas determined priority by national priorities
Post-Primary	**Children and youth acquire transferable, 21st-century skills for work, life, and future learning at the end of lower secondary**	Children and youth acquire proficient reading skills, demonstrate basic problem solving skills, and have an understanding of citizenship values	Children and youth demonstrate proficient skills for problem solving and critical thinking though academic content (typically in the areas of literacy, mathematics, and science)	Learning in other areas determined priority by national priorities

Early Childhood

Global Goal: Young children start school on time and are ready to learn

Domains: Physical well-being, social and emotional, literacy and communication (with a focus on oral language development), learning approaches and cognition, numeracy and mathematics.

Minimum goal indicator: (1) Children under five are free from chronic nutritional deficits, developmental difficulties, and health conditions; (2) NIR into first grade of primary.

Rationale: A healthy life is a fundamental human right; moreover, children who suffer from health and nutritional problems or developmental difficulties early in life demonstrate lower learning outcomes.[48] Anthropomorphic indicators are measured in most countries and developmental assessments are used within and across some countries on a population basis.[49] Outcomes related to health and nutrition can be viewed as "outcomes for learning" rather than "learning outcomes," but given the profound impact these variables have on children's learning development, they are considered in this indicator. The learning experience provided through primary education should start at the right moment in a child's life; therefore, it is important that education systems ensure that children are enrolled in primary at the optimal entry age.

Potential targets and measures:
- **% of children enrolling in primary school on-time.** Data source: UIS data on school entry (NIR).

- **% of children under 5 years not suffering from stunting (low height-for-age).**

- **% of children under 5 years not underweight (low weight-for-age).**

Example of data source: UNICEF-WHO-World Bank Joint Child Malnutrition Estimates (114 countries with data since 2005)

- **% of children ages 3-5 considered developmentally on track.** Examples of data sources: UNICEF MICS4-Early Child Development Index (55 countries);[50] Various developmental assessments.[51]

- **% of countries meeting minimum threshold on ECCE index.** Threshold would be defined in consultation with countries, technical experts and task force. Example of data source: *EFA Global Monitoring Report* ECCE Index (68 countries)[52]

Desirable goal indicator: Children demonstrate competence in domains associated with school readiness, including physical well-being, social and emotional, language, numeracy, learning approaches and cognition.

As a minimum guideline, learning outcomes in the following domains should be measured: physical well-being, social and emotional, language and communication, numeracy and mathematics, and learning approaches and cognition.

Rationale: Learning outcomes in these five domains are considered important by researchers, practitioners, and policymakers in the field of early childhood care and education.[53] They are also predictive of later school achievement and cognitive development.[54] Assessments that measure school readiness along these domains are increasingly being developed and used across populations and across countries.[55] School readiness is considered a desirable goal because not all countries have developed early learning standards, pre-primary enrollment rates are low especially in low-income countries,[56] and the quality of early childhood programs is variable.

Potential targets and measures:

- **% of children ready to learn at school upon school entry.**
  Examples of data sources: Early Development Instrument (EDI) (24 countries, though not all administer at the population level); School Readiness Instrument (India); East-Asia Pacific Early Child Development Scales (7 countries).

- **% of children receiving adequate scores on scales of learning in specific domains**
  Examples of data sources: many, including Strengths and Difficulties Questionnaire, Adaptive Behaviour Scale (social and emotional development); Peabody Picture Vocabulary Test (receptive language subdomain of literacy and communication), etc.[57]

National Goals

Outcomes related to two additional learning domains— "culture and the arts" and "science and technology" are included in some country-level frameworks[58] but are currently not globally emphasized measures of learning in early childhood. Therefore, the group recommends that these domains be considered for measurement at a national or local level and not as part of a global goal for early childhood for the time being. However there are several examples of how these goals are measured:

In South Korea, the "Nuri-Curriculum-Based Early Childhood Assessment Scale" assesses five-year-old children's knowledge, skills and attitudes in areas determined by the national curriculum. This measure includes teacher ratings on general development, including in the area of "inquiry of nature" and "arts."

"Plan d'étude Roman" in Switzerland assesses children's knowledge against curricular content. In science and technology, observable indicators include:

- Introduction to technology—child uses simple tools,

while being aware of their dangers (scissors, candles, lighters, electrical appliances)

- Introduction to technology—child is initiated to use of computers, and uses the learning resources available (CD-ROM, school website)

- Inquiry of nature—child distinguishes seasons by observing variations in temperatures

Primary

Global Goal: Children acquire basic literacy, numeracy, and global citizenship skills by the end of primary school.

Domains: Literacy and communication, numeracy and mathematics, social and emotional, with some elements of culture and the arts, science and technology, and learning approaches and cognition.

Minimum goal indicator: All children complete primary school and acquire basic reading skills by the age of 15.

Rationale: By setting the limit age for primary school completion at 15, children who start school late or repeat grades can be included in the count of children completing primary school. In 34 countries, more than 15% of primary school students are older than the intended age group; in countries such as Guinea-Bissau, Nepal, and Equatorial Guinea, more than 30% of children are older than the intended age.[59]

Reading skills in early primary can be predictive of a child's ability to read 5 or even 10 years later.[60] This is largely due to the fact that reading skills are self-reinforcing—children who are strong readers read more and encounter more novel words. Through wide reading, children develop a larger vocabulary, which in turn helps them read and understand new material.[61]

Across languages, learning to read follows a similar trajectory,[62] although what varies is the length of time needed to acquire basic literacy skills such as sound/symbol relationships. Approximately 90 countries participate in assessments of reading that are expected to be comparable across countries as early as 2017.[63]

Potential targets and measures:
- **% of children completing primary school by age 15.**
  Example of data source: Percentage of population of age 15 who have completed at least primary education (population census or household surveys)

- **% of children scoring at least a "basic" international benchmarked level for reading.** The definition of "basic" would be developed by the Measures and Methods Working Group based on an analysis of existing validated measures of reading. Example of data sources: Population-based assessments, e.g., Uwezo, Annual Status of Education Report (ASER); school-based assessments where appropriate, e.g., Progress in International Reading Literacy Study (PIRLS)/pre-PIRLS, Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), Analysis Programme of the CONFEMEN Education Systems (PASEC), Latin America Laboratory for Assessment (LLECE), and national assessments. (Note: Internationally comparable data would be available for more than 90 countries, but not all countries covered by these data sources.)

Desirable goal: All children complete primary school on time (at the expected graduation age) and acquire intermediate skills in literacy and communication, numeracy and mathematics, and global citizenship (including social and cultural skills, multilingual skills, higher-order skills, etc.).

Rationale: In addition to literacy, numeracy is seen as a foundational skill for learning and participation in society.[64] Lack of numeracy skills leads to many challenges, such as difficulty with everyday tasks and a lack of opportunities for the future.[65] Literacy and numeracy are widely measured in primary school using assessments that are standardized and used across populations in multiple countries (see below).

There is an increasing emphasis on global citizenship as a goal of education and learning.[66] The UN secretary-general's Education First initiative defines global citizenship as skills that help people "forge more just, peaceful, tolerant and inclusive societies" and give them the "understanding, skills and values they need to cooperate in resolving the interconnected challenges of the 21st century."[67]

Potential targets and measures:
- **% of children completing primary school at the expected graduation age.**
  Example of data source: Percentage of population who have completed at least primary education at the expected age (population census or household surveys)

- **% of students scoring at least at "proficient" international benchmarked level for reading.** The definition of "proficient" would be developed by the Measures and Methods Working Group based on an analysis of existing validated measures of reading. Examples of data sources: school-based assessments, e.g., PIRLS/pre-PIRLS, SACMEQ, PASEC, LLECE, national assessments when no regional/international assessments available.

- **% of students scoring at least at "proficient" international benchmarked level for mathematics.** The definition of "proficient" would be developed by the Measures and Methods Working Group based on an analysis of existing validated measures of reading. Examples of data sources: school-based assessments, e.g., Trends in Mathematics and Science Study (TIMSS), SACMEQ, PASEC, LLECE, national

assessments when no regional/international assessments available.

- **% of children receiving "proficient" scores on assessment related to global citizenship** (typically content considered part of the social and emotional domain, but could also be found in culture and the arts, learning approaches and cognition, and science and technology) as chosen by the country. Potential data sources: For primary-age students, no international assessments of global citizenship are currently used. However, some assessments are emerging at the national and regional level. For example,

  - Uwezo in Eastern Africa is a household survey that includes a bonus question that covers social and cultural knowledge that may not necessarily be acquired by enrolling in school. Some of the issues that have been covered over time are ethno-mathematics (mathematics related to the culture), parts of the body, and the flag as a symbol of national unity. Children may answer this question in any language including mother tongue. The Developmental Asset Lists contain items related to identity, social justice, caring for others, responsibility, social competence, etc. This measure has been translated in to 15 languages and has competency lists spanning from early childhood through post-primary.

  - Various countries, including Chile, Peru, Guatemala, South Korea, India, the United States and Colombia have developed country-specific assessments of social and civic learning.[68]

National

Progress toward some learning outcomes may be better assessed at the country level. Some pre-reading skills, such as decoding letters and words, are very important to later reading ability, yet vary in the time it takes children to master these skills. Measuring early pre-reading skills is also important because it allows countries to react quickly in order to address potential problems while children are still in schools.

Examples of assessments that could be used at the national level to measure incremental progress toward goals in literacy and numeracy include:

- Early Grade Reading Assessment (EGRA), 44 countries

- Literacy Boost, 15 countries

- Early Grade Writing Assessment (EGWA), in pilot phase

- Early Grade Math Assessment (EGMA), 11 countries

Other domains are closely related to national priorities, needs, cultures and economy. Examples of how some additional domains may be tracked at a national level include:

Physical well-being

- Literacy Boost Child Questionnaire (Burundi, Ethiopia, Malawi, Philippines, South Africa, Zimbabwe)

- Korean Children and Youth Panel Survey 2010 (KCYPS) and The Research on the Actual Status and the Quality in School Education in 2011 (Korea)

- Physical and Social well-being framework (India)

- Knowledge, Attitudes and Practices (KAPs Surveys) on thematic issues related to health (various countries worldwide).

Culture and the arts

- 2011 End-of-Year SA Assessment Tests and Objectives Written by Faculty 2nd–6th Grade Ixil (Guatemala)

- 2009 Survey on Korean Children and Youth's Activities and Culture ; Korean Children and Youth Panel Survey 2010 (KCYPS); the Research on the

Actual Status and the Quality in School Education in 2011 (South Korea)

- National Assessment of Educational Progress (NAEP) (United States)

Learning approaches and cognition

- Evaluación de habilidades para el aprendizaje and Evaluando aprendizajes ¿Qué, cómo y para qué? (Peru)

- KEDI Comprehensive Test and The Research on the Actual Status and the Quality in School Education in 2011 (South Korea)

Science and Technology

- National Assessment of Educational Achievement (NAEA) and KEDI Comprehensive Test (South Korea)

- Learning Guarantee Program (India)

- National Assessment of Educational Progress (NAEP), (United States)

Postprimary

Global Goal: Children and youth acquire transferable, 21st-century skills for work, life and future learning.

Domains: Social and emotional, literacy and communication, learning approaches and cognition, numeracy and mathematics, science and technology, with some subdomains of culture and the arts and physical wellbeing.

Minimum goal indicator: Children and youth demonstrate proficient reading skills, basic problem solving skills, and develop citizenship values.

Rationale: Skilled adolescent and adult readers and writers are far more likely to be successful at home and

in the workplace than their unskilled peers.[69] There is an increasing emphasis on education for citizenship at the global level since it is related to promote mutual understanding and intercultural dialogue as stated in the Universal Declaration of Human Rights (1948).

Potential targets and measures:
- **% of children and youth scoring at least at "proficient" international benchmarked level for reading.** This target is the same as the desirable target for primary. The definition of "proficient" would be developed by the Measures and Methods Working Group based on an analysis of existing validated measures of reading. Examples of data sources: Examples of data sources: Programme for International Student Assessment (PISA) (75 countries), Programme for the International Assessment of Adult Competencies (PIAAC) (23 countries), Literacy Assessment and Monitoring Programme (LAMP) (12 countries)

- **% of children and youth scoring at least at "basic" international benchmarked level in problem solving.** The definition of "basic" would be developed by the Measures and Methods Working Group based on an analysis of existing validated measures of problem solving. Examples of data sources: Programme for International Student Assessment (PISA) (75 countries), Programme for the International Assessment of Adult Competencies (PIAAC) (23 countries).

- **% of children and youth receiving adequate scores on global citizenship assessment.** Example of data source: IEA International Civic and Citizenship Education Study (ICCS) (38 countries)[70]

Desirable goal indicator: Children and youth demonstrate proficient skills for problem solving and critical thinking though academic content (typically in the areas of literacy, mathematics and science) by the end of lower secondary.

Rationale: Examination of the higher order skills of experienced decision makers points to flexibility and criti-

cal thinking skills as key components of the procedures they apply.[71] Occupations are becoming increasingly reliant on a cognitively flexible and adaptable work-force.[72] Competencies in these areas are currently being measured in more than 70 economies worldwide.[73] However, this goal is desirable and not minimum because many of the most important 21st-century skills are difficult to measure in a standardized way, especially independent of academic content for those who have not attended secondary school.

Potential targets and measures:
- **% of children and youth scoring at least at "proficient" international benchmarked level for critical thinking and problem solving as demonstrated on tasks related to reading, mathematics, and scientific literacy.** The definition of "proficient" would be developed by the Measures and Methods Working Group based on an analysis of existing validated measures of learning in these areas. Examples of data sources: School-based assessments used in secondary, such as PISA, TIMSS, and national assessments.

- **% of children demonstrating competencies in 21st century skills.**
  Examples of data sources: The Assessment & Teaching of 21st Century Skills (ATC21S) project examined multiple frameworks for 21st-century skills and developed a framework for assessment in two skill areas: collaborative problem-solving and ICT digital literacy (ATC21S 2012).

National
Most of the currently administered international assessments for postprimary focus on academic content in literacy, science and mathematics in secondary schooling; nevertheless, some countries have ongoing efforts to measure other domains nationally, such as "Culture and the Arts" and "Science and Technology." For example:

- National Citizenship Competencies, Colombia, is a school-based assessment that is administered to students in grades 5 and 9, in social awareness, leadership, viewing on one's self, resilience, moral and ethics, values, and conflict resolution.

- National Assessment Programme on Civics and Citizenship, Australia, is a school-based assessment that is administered to students in years 6–10 and examines civic knowledge and understanding the skills and values needed to active citizenship.[74]

- In 2012 Arnhen Cito, the Netherlands, started Balance of the social outcomes in primary education, a first attempt to measure social competencies, civic, moral and ethical values to 11-year-olds (at the end of their primary education)

These are just several of many examples of national assessments in post-primary that measure progress toward learning goals prioritized at the country level.

Next Steps

This prototype framework, along with a toolkit to guide consultations, will be available on the Learning Metrics Task Force Web site (www.brookings.edu/learningmetrics) for consultation and comments by all interested stakeholders. Translations will be available in Arabic, French, English, Russian, and Spanish.

Feedback should be submitted to learningmetrics@brookings.edu by January 18, 2013. All feedback received by this date will be organized and considered by the Measures and Methods Working Group as it works rapidly to revise the framework based on consultation results.

Once the proposed goals and indicators have been determined, the working group will conduct a deeper analysis of the measures and targets that may be used to measure learning toward these goals. The working group will then propose a model (or several models if no consensus is reached) to present to the task force at the in-person meeting on February 20–21, 2013,

in Dubai. The task force will deliberate and make changes to the framework during the meeting, and the final report on measuring learning outcomes will be available in April 2013.

Discussion questions for use with this document:

a. Do you think the overall framework structure—with global minimum, global desirable, and nationally defined goals—is a good way to capture learning at the global and national levels? Do you think these terms accurately describe the different types of goals? If not, what terms would you recommend instead?

b. Consider the global goals proposed in this prototype. Do these goals capture the major learning outcomes that should be expected at the early childhood, primary, and postprimary stages, respectively? If not, how would you change these goals?

c. Consider the global minimum indicators. Do these describe foundational outcomes that support (a) later stages of learning and (b) more advanced levels of learning? What indicators, if any, should be modified? What indicators are missing? What additional evidence would you need to decide?

d. Consider the potential targets and measures for the global minimum indicators. What additional measures are you aware of that could capture learning toward these targets? Do you have any concerns about the potential measures? How important is it to your work that these measures be internationally or regionally comparable (i.e., providing information on how the country or countries where you work measure up to other countries in the region or world)?

e. Consider the global desirable indicators. Do these describe outcomes that are (a) necessary for accessing further educational opportunities and (b) based on skills needed in the next 15 to 20 years and beyond? What indicators, if any, should be modified? What indicators are missing?

f. Consider the potential targets and measures for the global desirable indicators. What additional measures are you aware of that could capture learning toward these targets? Do you have any concerns about the potential measures? How important is it to your work that these measures be internationally or regionally comparable (i.e., providing information on how the country or countries where you work measure up to other countries in the region or world)?

g. Consider the recommendations for national indicators. Are these areas of learning outcomes best measured at a country or local level? What additional examples of assessments are you aware of that have been useful to measure progress toward learning goals in areas such as physical well-being, social and emotional , culture and the arts, learning approaches and cognition, and science and technology? Should all countries measure progress toward unique national goals?

If you are recommending additional learning measures, please include the following information in the consultation notes: name of measure, author, country or countries where used, domains covered, age group or grade levels covered, and website (if available).

Please submit responses to learningmetrics@brookings.edu.

Endnotes

1. These data should be interpreted with caution. Due to drastic expansion of primary schooling between 2000 and 2007, the children previously excluded from the primary school population were included in the sample for SACMEQ III. As a result, the true performance of Mozambique is likely to be lower than was reflected in SACMEQ II (Spaull 2011).)

2. It needs to be kept in mind that this is relative to the degree that the measurement is a proxy for future capabilities.

3. For details on current efforts in monitoring access to schooling, see the UIS web site, www.uis.unesco.org.

4. One of the principal factors affecting the accuracy of enrollment ratios is the quality of population data. Population data of poor quality can in turn generate enrollment data of poor quality.

5. Multiple terms have been used to describe development and learning at this stage: school readiness, early learning and/or development, readiness to learn in school, etc. Because these terms have different connotations in different countries and languages, the working group recommended that the LMTF not endorse one term but rather focus on describing the goals for this stage.

6. Data were obtained from EGRA Tracker (2013), https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=188, which includes a summary and description of the purpose, scope and sample size of assessments of foundational reading skills in developing countries. Note that these are low-stakes assessments and are not necessarily representative of the whole population—they also include assessments designed for program design and evaluation, which are generally administered to a small sample.

7. "Phonological awareness" and "sound-print correspondence" are the terms used for alphabetic languages. However, oral language development skills such as sound awareness and awareness and skill with sound-print correspondence have also been developed for logographic languages.

8. Between all the citizen led assessments in South Asia, East Africa and West Africa, close to a million children are assessed each year.

9. The countries are Indonesia, Uganda, Zimbabwe, South Africa, Ethiopia, Malawi, Philippines, Burundi, Malawi, Mozambique, Mali, Bangladesh, Nepal, Guatemala, Pakistan, Vietnam, Haiti, Yemen and Egypt.

10. One has to assume in this case that children understood the tasks presented to them.

11. This would, however, exclude the languages where the notion of a letter or word does not exist.

12. The ASER instrument is accompanied by a background questionnaire that records the child's age, grade-level and other demographic characteristics.

13. LAMP is a household-based assessment of adults' (15 years or older) reading and numeracy skills. A report will be released in 2013 outlining the results from four countries in various regions (Jordan, Mongolia, Palestine and Paraguay). The instruments and procedures have also been field-tested in five other countries (Afghanistan, El Salvador, Morocco, Niger and Vietnam). PIAAC is a household-based survey of adults age 16–64 focuses

on cognitive and workplace skills needed for successful participation in 21st-century society and the global economy. Specifically, PIAAC measures relationships between individuals' educational background, workplace experiences and skills, occupational attainment, use of information and communications technology, and cognitive skills in the areas of literacy, numeracy and problem solving. A report will be released at the end of 2013 outlining the results from 24 participating countries.

14. For instance, while exploring the associations among a given variable (i.e., performance) and a set of covariates, lack of variability in a measure reflecting floor effects may lead to artificially low or nonexistent associations.

15. For more details on the notion of international statistics as a public good, see the section below titled "Information as a Global Public Good."

16. The Democratic Republic of Congo, Dominican Republic, Iraq, Jordan, Kenya, Liberia, Malawi, Mali, Morocco, Nicaragua, Rwanda, Zambia.

17. Optionally, TIMSS can also be administered in grades 5 or 6, instead of in grade 4.

18. It is important to note that the largest share of any measurement endeavor is associated with the in-country operations, particularly with the data generation activities. Thus, while some developmental costs (in country or covered through "international fees" when participating in an international study) are usually easy to report and, therefore, are clearly visible, they do not represent a large share of the total costs. At the same time, operational costs are directly associated with the interest in being able to report according to different information breakdowns; the more detailed information is sought, the larger the sample size, and therefore, the more demanding the field operations.

19. A "construct" is the conceptual entity that is the focus of the measurement effort. It is an operationalized concept that usually corresponds to a latent (no directly observable) trait. See Cronbach and Meehl (1955).

20. The International Test Commission has done extensive work that provides the world standards in this area. See http://www.intestcom.org/guidelines/index.php.

21. A "metric" is a particular way (operationalization, procedures, standards and/or tools) of measuring a given construct.

22. In this respect, a rapidly expanding literature on the fundamental importance of indigenous knowledge and traditional learning has already been summarized by Nakashima (2010) and the discussion is ongoing in UNESCO's Local and Indigenous Knowledge Systems Program (LINKS).

23. The 2011 review of the International Standard Classification of Education (UNESCO/UIS 2012) provides a comprehensive account of what is to be understood as formal and nonformal learning (paragraphs 35–44). While both are intentional in character and usually organized in institutional settings, the latter does not necessarily belong to a specific and fully structured pathway that leads to further learning or does not lead to certifications that are recognized by the relevant authorities.

24. This includes Kirsch and Mosenthal, LAMP, and PIAAC/OECD.

25. The synthesis is based on analysis of online education e-discussions with civil society, engaging more than 19,000 people from over 100 countries; the ongoing EFA regional meetings and the Collective Consultation of NGOs on EFA; a member states briefing involving over 40 governments; a consultation with representatives from the private sector and from donor agencies; and the global education meeting held in Dakar earlier this year, involving over 100 including member states, youth groups, the private sector, civil society organizations and UN agencies.

26. Assessments that are intended to yield equivalent

information across different subpopulations at the national, regional or global level.

27. The full list of Measures and Methods Working Group members is available at the beginning of this report

28. "On time" equals the officially expected graduation age. To measure it, a proper record of ages is needed.

29. Or any other relevant age-group, depending on the official graduation age and the level of existing late graduation.

30. Or any other relevant age-group, depending on the official graduation age and the level of existing late graduation.

31. UNESCO (2012).

32. For a list of task force members, see http://www.brookings.edu/~/media/Centers/universal%20education/learning%20metrics%20task%20force/LMTF_Members_Oct2012.pdf

33. For a list of Working Group members, see http://www.brookings.edu/about/centers/universal-education/learning-metrics-task-force/working-groups.

34. A prototype is an early model developed to test a concept or design with stakeholders in order to improve upon it. This methodology is most often used in engineering and software design, but is increasingly being used in the social sciences. See Coughlan, Suri and Canales (2007).

35. See http://www.uis.unesco.org/StatisticalCapacityBuilding/Workshop%20Documents/Education%20workshop%20dox/2010%20ISCED%20TAP%20IV%20Montreal/NFE_CLA_Eurostat_EN.pdf.

36. Ibid.

37. Ibid.

38. United Nations, "The Universal Declaration of Human Rights" (1948), Article 26.

39. UIS (2012). Out of 200 countries with data, primary education is compulsory in 190.

40. UNESCO (2012).

41. UIS (2012).

42. Ibid.

43. Ibid.

44. UNESCO (1997).

45. See Shonkoff and Phillips (2000).

46. Internationally and regionally comparable learning assessments can be useful tools for measuring progress toward global goals because they measure the same constructs using an equivalent approach. However, the coverage of these assessments does not extend to many low-income countries and small states for a variety of reasons, including a limited demand for comparing learning outcomes with other countries, sufficient country-developed learning assessments, and limited capacity and resources for participation. The Measures and Methods Working Group will investigate ways to use data from various approaches to assessing learning. The working group is aware that concepts such as "basic" and "proficient" do not mean the same level of ability across measures.

47. Even if there are tools available to measure learning in a given area, it cannot be assumed that there is capacity at a country level to collect, collate and analyze data. The working group will explore capacity in these areas in subsequent reports.

48. Grantham-McGregor et al. (2007); Paxson and Schady (2007).

49. For example, the UNICEF Multiple Indicator Cluster Survey (MICS).

50. For more information on the multicountry assessments listed throughout this document, please see Simons (2013).

51. For a comprehensive review of measures, see Fernald et al. (2007).

52. UNESCO (2012). The ECCE Index is based on the mean of three indicators: percentage of children surviving beyond fifth birthday; percentage of children under five not suffering from stunting; and percentage of children age three to seven enrolled in pre-primary or primary education. The education component of this index could be used to measure progress toward the desirable goal.

53. In a review of common constructs in Early Learning and Development Standards (ELDS) in 12 countries, four domains emerged as being part of each individual country's standards: cognitive development, language and literacy, social and emotional, and health and motor. See Kagan and Britto (2007).

54. See, for example, Romano et al. (2010); Epstein (2009); Parker and Asher (1987); Raver et al. (2005); Alexander, Entwisle and Dauber (1993); McClelland, Morrison and Holmes (2000); Yen, Konold and McDermott (2004); Howse et al. (2003); and Hart and Risley (2003).

55. E.g., Rao et al. (2012).

56. UNESCO (2012).

57. See Fernald et al. (2007).

58. Kagan and Britto (2007).

59. UIS (2012).

60. McCardle, Scarborough and Catts (2001); Cunningham and Stanovich (1997); Juel (1988).

61. Aga Khan Foundation (2010).

62. Abdazi (2012).

63. The Global Partnership for Education (GPE) is working with partners to organize an initiative that will allow for results from four reading assessments to be compared: LLECE, PASEC, PIRLS/PrePIRLS, and SACMEQ. LLECE's current cycle (called TERCE) will involve 14 countries; PASEC comprises 21 countries; out of those 21 countries, 10 will participate in the next round, in 2015. The latest edition of PIRLS/ prePIRLS involved a total of 49 countries; SACMEQ involves 15 countries.

64. Orrill (2001).

65. Kilpatrick, Swafford, and Findell (2001).

66. UNSG (2012).

67. Ibid.

68. LMTF Measures and Methods Working Group assessments canvassing, ongoing.

69. Graham and Perin (2007); Sum et al. (2007).

70. Austria, Belgium (Flemish), Bulgaria, Chile, Chinese Taipei, Colombia, Cyprus, Czech Republic, Denmark, Dominican Republic, England, Estonia, Finland, Greece, Guatemala, Hong Kong SAR, Indonesia, Ireland, Italy, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Netherlands, New Zealand, Norway, Paraguay, Poland, Russian Federation, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, and Thailand.

71. Helsdingen, Van Gog, and Van Merriënboer (2009).

72. Billett (1998); Smith (2003).

73. E.g. PISA, PIAAC.

74. According to ISCED classification, this would be once in ISCED 1 and during the first year of ISCED 2.

References

Abdazi, H. 2012. *Developing Cross-Language Metrics for Reading Fluency Measurement: Some Issues and Options*. Washington: Global Partnership for Education.

Aga Khan Foundation. 2010. *Improving Learning Achievement in Early Primary in Low-Income Countries: A Review of the Research*. Geneva: Aga Khan Foundation.

Alexander, K. L., D. R. Entwisle, and S. L. Dauber. 1993. First-Grade Classroom Behavior: Its Short-and Long-Term Consequences for School Performance. *Child Development* 64 (3): 801–14.

ASER Centre. 2012. *ASER Center: Evidence for Action*. http://www.asercentre.org/

Ary, D., L. Jacobs, C. Sorensen, and A. Razavieh. 2006. *Introduction to Research in Education*, 8th ed. Wadsworth CENGAGE Learning USA.

ATC21S. 2012. ATC21S Project Overview, July 2012. http://atc21s.org/wp-content/uploads/2012/07/CiscoGlobalEdATC21SAmsterdamBrochure_D15_V7-1.pdf.

AusAID. 2012. *Australia's Comprehensive Aid Policy Framework*. http://www.ausaid.gov.au/makediff/pages/capf.aspx/

Billett, S. 1998. Appropriation and Ontogeny: Identifying Compatibility between Cognitive and Sociocultural Contributions to Adult Learning and Development. *International Journal of Lifelong Education* 17: 21–34.

Beatty, A., and L. Pritchett. 2012. From *Schooling Goals to Learning Goals: How Fast Can Student Learning Improve?* CDG Policy Paper 012. Washington: Center for Global Development.

Clarke, M. 2012. *What Matters Most for Student Assessment Systems: A Framework Paper.* Washington: World Bank.

Coughlan, P., J. F. Suri, and K. Canales. 2007. Prototypes as (Design) Tools for Behavioral and Organizational Change. *Journal of Applied Behavioral Science* 43 (1): 1–13. http://www.ideo.com/images/uploads/news/pdfs/Prototypes_as_Design_Tools_1.pdf.

Cronbach, L. J., and P. E. Meehl. 1955. Construct Validity in Psychological Tests. *Psychological Bulletin* 52: 281–302.

Cunningham, A. E., and K. E. Stanovich. 1997. Early Reading Acquisition and Its Relation to Reading Experience and Ability 10 Years Later. *Developmental Psychology* 33 (6): 934.

Duncan, G., C. Dowsett, A. Claessens, K. Magnuson, A. Huston, P. Klebanov, L. Pagani., L. Feinstein, M. Engel, J. Brooks-Gunn, H. Sexton, K. Duckworth, and C. Japel. 2007. School Readiness and Later Achievement. *Developmental Psychology* 43 (6): 1428–46.

Epstein, A. 2009. *Me, You, Us: Social-Emotional Learning in Preschool*. Ypsilanti, Mich.: HighScope.

European Commission. 2006a. *Classification of Learning Activities-Manual*. Luxembourg: European Communities.

———. 2006b. *European Parliament and The Key Competences for Lifelong Learning: A European Framework*. Brussels: European Communities.

Fernald, L. C. H., P. Kariger, P. Engle and A. Raikes. 2007. Examining Early Childhood Development in Low-Income Countries: A Toolkit for the Assessment of Children in the First Five Years of Life. Washington: World Bank. http://siteresources.worldbank.org/INTCY/Resources/395766-1187899515414/Examining_ECD_Toolkit_FULL.pdf.

GPE (Global Partnership for Education). 2012. *Strategic Plan* 2012–2015. http://www.globalpartnership.org/media/docs/library/GPE_Stategic_Plan_2012-2015_English.pdf

Graham, S., and D. Perin. 2007. A Meta-Analysis of Writing Instruction for Adolescent Students. *Journal of Educational Psychology* 99 (3): 445.

Grantham-McGregor, S., Y. B. Cheung, S. Cueto, P. Glewwe, L. Richter, B. Strupp, and International Child Development Steering Group. 2007. Child Development in Developing Countries 1: Developmental Potential in the First 5 Years for Children in Developing Countries. *The Lancet* 369 (9555): 60–71.

Hambleton, R. 2005. Issues, Designs and Technical Guidelines for Adapting Tests into Multiple Languages and Cultures. In *Adapting Psychological and Educational Tests for Cross-Cultural Assessment*, edited by R. Hambleton, P. Merenda, and C. Spielberger. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hart, B., and T. R. Risley. 2003. The Early Catastrophe: The 30 Million Word Gap by Age 3. *American Educator* 27 (1): 4–9.

Helsdingen, A. S., T. Van Gog, and J. J. G. Van Merriënboer. 2009. Critical Thinking Instruction and Contextual Interference to Increase Cognitive Flexibility in Complex Judgment. Paper presented at the Joint Meeting of the Scientific Network on "Developing Critical and Flexible Thinking" and the European Network on Epistemological beliefs. June, 3–5, 2009, Marche-en-Famenne, Belgium.

Hutchison, D. and Schagen, I. 2006. "Comparisons between PISA and TIMSS – Are We the Man with Two Watches?" National Foundation for Educational Research.

Hockenbury, D. H., and S. E. Hockenbury. 2010. *Psychology*, 5th Ed. New York: Worth Publishers.

Howse, R. B., G. Lange, D. C. Farran, and C. D. Boyles. 2003. Motivation and Self-Regulation as Predictors of Achievement in Economically Disadvantaged Young Children. *Journal of Experimental Education* 71 (2): 151–74.

IEA (International Assessment for the Evaluation of Educational Achievement). 2011. Trends in Mathematics and Science Study 2011. http://www.iea.nt/timss_2011.html.

Janus, M., and D. Offord. 2007. Development and Psychometric Properties of the Early Development Instrument (EDI): A Measure of Children's School Readiness. *Canadian Journal of Behavioral Science* 39 (1): 1–22.

Juel, C. 1988. Learning to Read and Write: A Longitudinal Study of 54 Children from First through Fourth Grades. *Journal of Educational Psychology* 80 (4): 437.

Kagan, S. L., and P. R. Britto. 2007. *Frequently Occurring Constructs: What Countries Want Their Young Children to Know and Be Able to Do*. New York: United Nations.

Kellaghan, T., V. Greaney, and S. Murray. 2009. *Using the Results of a National Assessment of Educational Achievement*. National Assessments of Educational Achievement Series, vol. 5. Washington: World Bank.

Kilpatrick, J., J. Swafford, and B. Findell. 2001. *Adding It Up: Helping Children Learn Mathematics*. Washington: National Academies Press.

Kirsch, I. S., and P. B. Mosenthal. 1990. Exploring Document Literacy: Variables Underlying the Performance of Young Adults. *Reading Research Quarterly* 25: 5–30.

Lewin, K. M. 2011. *Making Rights Realities: Researching Educational Access, Transitions and Equity*. CREATE Synthetic Report. Centre for International Education, University of Sussex.

LMTF (Learning Metrics Task Force). 2013. *Toward Universal Learning: What Every Child Should Learn. Report No. 1 of the Learning Metrics Task Force*. Montreal and Washington: UNESCO Institute for Statistics and Center for Universal Education at the Brookings Institution. http://www.brookings.edu/about/centers/universal-education/learning-metrics-task-force/~/media/56D69BF9960F4442864F28AE28983248.ashx

McCardle, P., H. S. Scarborough, and H. W. Catts. 2001. Predicting, Explaining, and Preventing Children's Reading Difficulties. *Learning Disabilities Research & Practice* 16 (4): 230–39.

McClelland, M. M., F. J. Morrison, and D. L. Holmes. 2000. Children at Risk for Early Academic Problems: The Role of Learning-Related Social Skills. *Early Childhood Research Quarterly* 15 (3): 307–29.

Motala, S., V. Dieltiens, and Y. Sayed. 2009. Physical Access to Schooling in South Africa: Mapping Dropout, Repetition and Age-Grade Progression in Two Districts. *Comparative Education* 45 (2): 251–63.

Mullis, I., M. Martin, A. Kennedy, K. Trong, and M. Sainsbury. 2009. PIRLS 2011 Assessment Framework. TIMSS and PIRLS International Study Center, Lunch School of Education, Boston College.

Nakashima, Douglas, ed. 2010. *Indigenous Knowledge in Global Policies and Practice for Education, Science and Culture*. Paris: UNESCO.

OECD (Organization for Economic Cooperation and Development). 2009. *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science*. Program for International Student Assessment. Paris: OECD.

Orrill, R. 2001. *Mathematics, Numeracy and Democracy. In Mathematics and Democracy: The Case for Quantitative Literacy*, edited by L. A. Steen. Princeton, NJ: National Council on Education and the Disciplines.

Parker, J. G., and S. R. Asher. 1987. Peer Relations and Later Personal Adjustment: Are Low-Accepted Children at Risk? *Psychological Bulletin* 102 (3): 357–89.

Paxson, C., and N. Schady. 2007. Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health, and Parenting. *Journal of Human Resources* 42 (1): 49–84.

Postlehwaite, T. N. 2004. *Monitoring Educational Achievement*. Fundamental of Educational Planning Series. Paris: UNESCO/IIEP.

Rao, N., P. L. Engle, J. Sun and M. Ng. 2012. *Development of the East Asia-Pacific Early Child Development Scales (EAP-ECDS) Report of the Consultancy (Stage II)*. Hong Kong: University of Hong Kong.

Raver, C. C., R. Smith-Donald, T. Hayes, and S. M. Jones. 2005. Self-Regulation across Differing Risk and Sociocultural Contexts: Preliminary Findings from the Chicago School Readiness Project. In *Biennial Meeting of the Society for Research in Child Development, Atlanta*.

Romano, E., L. Babchishin, L. S. Pagani, and D. Kohen. 2010. School Readiness and Later Achievement: Replication and Extension Using a Nationwide Canadian Survey. *Developmental Psychology* 46: 995–1007.

RTI International. 2009. Early Grade Mathematics Assessment (EGMA): *A Conceptual Framework Based on Mathematics Skills Development in Children*. Washington: US Agency for International Development.

Save the Children. 2012. *Education Global Initiative: Moving Ahead on Education—A Focused Strategy for Achieving Our Education Goals* 2012–2015. London: Save the Children.

Schaffner, J. 2006. Measuring Literacy in Developing Country Household Surveys: Issues and Evidence. Background paper for EFA Global Monitoring Report.

Shonkoff, J., and D. Phillips. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington: National Academies Press.

Search Institute. 2006. *Developmental Assets List*. Minneapolis: Search Institute.

SG HLPEP (Secretary-General's High-Level Panel of Eminent Persons). 2013. *A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development: The Report of the High-Level Panel of Eminent Persons on the Post-2015 Development Agenda*. New York: United Nations. http://www.post2015hlp.org/wp-content/uploads/2013/05/UN-Report.pdf.

Shamblott, M. 2012. Introduction to the Work Sampling System for Head Start. Presentation to Pearson Clinical Assessment. http://psychcorp.pearson-assessments.com/hai/images/pdf/webinar/WorkSamplingSystemforHeadStartWebinarHand-out_5162012.pdf.

Simons, K. A. 2013. *Multi-Country Assessments of Learning*. Learning Metrics Task Force Discussion Document 1. Washington, DC: Brookings.

Smith, P. J. 2003. Workplace Learning and Flexible Delivery. *Review of Educational Research* 73: 53–88.

Spaull, N. 2011. *Primary School Performance in Botswana, Mozambique, Namibia and South Africa*. Working Paper 8. Paris: Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ).

Stephens, M., and M. Coleman. 2007. Comparing PIRLS and PISA with NAEP in Reading, Mathematics, and Science. Washington: US Department of Education.

Sum, A., Khatiwada, I., McLaughlin, J. Tobar, P., & Motroni, J. 2007. *An Assessment of the Labor Market, Income, Health, Social, Civic and Fiscal Consequences of Dropping Out of High School: Findings for Massachusetts Adults in the 21st Century*. Center for Labor Market Studies Northeastern University.

Tuck, Ron. 2007. *An Introductory Guide to National Qualifications Frameworks: Conceptual and Practical Issues for Policy Makers*. Geneva: International Labor Office.

UIS (UNESCO Institute for Statistics). N.d. *Glossary*. Montreal: UIS. http://glossary.uis.unesco.org/glossary/en/home.

———. 2009. The Next Generation of Literacy Statistics: Implementing the Literacy Assessment And Monitoring Programme (LAMP). Montreal: UIS.

———. 2010. *Database*. Montreal: UIS.

———.2012. *Global Education Digest 2012: Opportunities Lost—The Impact of Grade Repetition and Early School Leaving*. Montreal: UIS.

———. 2013. *UNESCO Institute for Statistics Database. January 2013 Release*. Montreal: UIS. http://stats.uis.unesco.org/.

UNESCO. 1997. I*nternational Standard Classification of Education.* Paris: UNESCO.

———. 2005. *Education for All: The Quality Imperative*. Paris: UNESCO.

———. 2012. *Education for All: Global Monitoring Report—Youth and Skills, Putting Education to Work*. Paris: UNESCO.

UNESCO/UIS. 2012. *International Standard Classification of Education: ISCED 2011*. Montreal: UNESCO/UIS.

UNESCO and UNICEF. 2013. *Envisioning Education in the Post-2015 Development Agenda: Thematic Consultation on Education in the Post-2015 Agenda*. http://www.worldwewant2015.org/education2015.

UNSG (Office of the UN Secretary-General). 2012. Education First. Our Priorities. http://globaleducationfirst.org/priorities.html.

USAID (US Agency for International Development). 2012. *2011 USAID Education Strategy Technical Notes*. http://pdf.usaid.gov/pdf_docs/PDACT681.pdf.

Vagh, S. B. 2012. Validating the ASER Testing Tools: Comparisons with Reading Fluency Measures and the Read India Measures. http://img.asercentre.org/docs/Aser%20survey/Tools%20validating_the_aser_testing_tools__oct_2012__2.pdf.

Wagner, D. A., A. Babson, and K. M. Murphy. 2011. How Much Is Learning Measurement Worth? Assessment Costs in Low-Income Countries. *Current Issues in Comparative Education* 14 (1): 8–11.

Wagner, D. A., Murphy, K. M., and H. De Korne. 2012. *Learning First: A Research Agenda for Improving Learning in Low-Income Countries*. Washington: Brookings Institution. http://www.brookings.edu/research/papers/2012/12/learning-first-wagner-murphy-de-korne.

Wise, S., and C. DeMars. 2005. Low Examinee Effort in Low-Stakes Assessment: Problems and Potential Solutions. *Educational Assessment* 10 (1): 1–17.

World Bank. 2011. *Learning for All: Investing in People's Knowledge and Skills to Promote Development—Education Strategy 2020*. Washington: World Bank.

Yen, C. J., T. R. Konold, and P. A. McDermott. 2004. Does Learning Behavior Augment Cognitive Ability as an Indicator of Academic Achievement? *Journal of School Psychology* 42 (2): 157–69.

UNESCO
INSTITUTE
for
STATISTICS

United Nations
Educational, Scientific and
Cultural Organization

B | Center for
Universal Education
at BROOKINGS

P.O. Box 6128, Succursale Centre-Ville
Montreal, Quebec H3C 3J7, Canada
+1 514 343 6880
www.uis.unesco.org

1775 Massachusetts Avenue, NW
Washington, DC 20036
202-797-6000
www.brookings.edu/universal-education