United Nations
Educational, Scientific and
Cultural Organization

International Institute
for Educational Planning

Learning at the bottom of the pyramid

Science, measurement, and policy in low-income countries

Edited by
Daniel A. Wagner,
Sharon Wolf, and
Robert F. Boruch

Learning at the bottom of the pyramid:
Science, measurement, and policy in low-income countries

Learning at the bottom of the pyramid:

Science, measurement, and policy in low-income countries

Edited by Daniel A. Wagner, Sharon Wolf, and Robert F. Boruch

United Nations
Educational, Scientific and
Cultural Organization

**International Institute
for Educational Planning**

The views and opinions expressed in this book are those of the authors and do not necessarily represent the views of UNESCO, IIEP, or the University of Pennsylvania. The designations employed and the presentation of material throughout this review do not imply the expression of any opinion whatsoever on the part of UNESCO or IIEP concerning the legal status of any country, territory, city or area or its authorities, or concerning its frontiers or boundaries.

The publication costs of this study have been covered by IIEP-UNESCO and the University of Pennsylvania.

Published by:
International Institute for Educational Planning
7–9 rue Eugène Delacroix, 75116 Paris, France
info@iiep.unesco.org
www.iiep.unesco.org

Cover design: Morgan A. Belveal and pngtree.com
Typesetting: Linéale Production
Printed in IIEP's printshop
ISBN: 978-92-803-1420-5

Preface

Goal 4 of the Sustainable Development Goals places a central emphasis on educational equity in its formulation: 'Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.' Reflecting on the progress made since 2000, when the first UN goals were established and the Education for All goals reaffirmed, we see tremendous expansion in educational access. But we also can identify serious shortcomings in the effort to ensure quality education for all children. Improving equity has been especially problematic for poor and marginalized children and youth in low-income countries. They continue to reside at the bottom of the pyramid.

This edited volume, *Learning at the Bottom of the Pyramid: Science, measurement, and policy in low-income countries*, collects the rich debates from the international conference held in Philadelphia on 2–3 March 2017. Some 50 invited experts contributed their perspectives and insights on how to better understand the science of learning in low-resourced settings worldwide, how learning can be measured in marginalized populations, and ways that new policy approaches can improve learning. These are matters of substantial concern to international agencies, foundations, policy-makers, education specialists, and the public at large. The complexity of the issues and the diverse backgrounds of participants ensured that the analyses and debates captured here have breadth and variety. It is clear that in order to achieve both inclusion and equity in improving the quality of education, a better understanding of learning in low-income societies should take a high priority.

Dan Wagner, UNESCO Chair and Professor of Education,
University of Pennsylvania

Suzanne Grant Lewis, Director, IIEP-UNESCO

Acknowledgements

We would like to thank the co-sponsors of the conference, including IIEP-UNESCO, UNESCO Unitwin, Ford Foundation, and Conrad N. Hilton Foundation, for their generous support. We also appreciate the support of the University of Pennsylvania in hosting the conference, along with its Graduate School of Education/International Educational Development Program, the International Literacy Institute, and Perry World House. For their helpful assistance, we thank Lauren Scicluna-Simon, Alysha Banerji, Morgan Belveal, Nathan Castillo, Danielle Combs, Erin D'Amelio, Ania Drabek, Emma James, Evan Jansen, Selene Lee, Anushka Mehta, Jessica Peng, Hyunjae Won, and Fatima Tuz Zahra.

Contents

List of figures, boxes, and tables

Figures

Boxes

Tables

List of abbreviations

ABL	Activity Based Learning
ACTRC	Assessment Curriculum and Technology Research Centre
AfL	Assessment for Learning
ALLS	Adult Literacy and Life Skills Survey
ASER	Annual Status of Education Report
BoP	bottom of the pyramid
DAC	Development Assistance Committee
ECCD	early childhood care and development
ECDI	early childhood care and development index
EFA	Education for All
ELM	Emergent Literacy and Math
EYE	Early Years Evaluation
EYE-TA	Early Years Evaluation – Teacher Assessment
GDP	gross domestic product
GNI	gross national income
IAS	international assessment studies
ICT	information and communications technology
IDELA	International Development and Early Learning Assessment
LAMP	Literary Assessment and Monitoring Programme
LICs	low-income countries
LLECE	Latin American Laboratory for Education Evaluation
LMICs	low- and middle-income countries
MDGs	UN Millenium Development Goals
NAEP	National Assessment of Educational Progress
NFE	non-formal education
ODA	Official Development Assistance
OECD	Organisation for Economic Co-operation and Development
PIAAC	Programme for the International Assessment of Adult Competencies
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SDGs	UN Sustainable Development Goals
SES	socio-economic status
SSA	sub-Saharan Africa
TIMSS	Trends in International Mathematics and Science Study
TVET	technical and vocational education and training
UIS	UNESCO Institute for Statistics

About the authors

Aaron Benavot served as Director of the UNESCO Global Education Monitoring Report from 2014 to 2017. He is currently Professor in the School of Education at Albany-State University of New York.

Penelope Bender was until 2017 a Senior Education Advisor on the USAID Basic Education team, and Lead for the implementation of Goal 1 of the USAID Education Strategy.

Robert F. Boruch is the University Trustee Chair Professor of Education and Statistics at the University of Pennsylvania.

Nicholas Burnett was founding Managing Director at Results for Development (R4D) in Washington DC, and is chair of the Board of UNESCO's International Institute for Educational Planning.

Andres S. Bustamante is a Postdoctoral Research Fellow at Temple University, Philadelphia.

Manuel Cardoso is an Education Specialist focused on Learning at UNICEF Headquarters in New York.

Esther Care is a Senior Fellow with the Brookings Institution, Washington DC, and Professorial Fellow at the Assessment Curriculum and Technology Research Centre at the University of Melbourne.

Amrita Chudgar is an Associate Professor of Education Policy at Michigan State University.

Luis Crouch is Chief Technical Officer in the International Development Group at the Research Triangle Institute.

Amy Jo Dowd is Senior Director for Education Research in Save the Children's Department of Education and Child Protection.

David K. Evans is a Senior Economist in the Chief Economist's Office for the Africa Region of the World Bank.

Michael G. Fast is a Managing Researcher at the American Institutes for Research (Washington, DC) and Director of the Education Standards and Assessment practice area.

Marlene Ferido is Chair of the Chemistry Education Group at the University of the Philippines-Diliman, and Philippine Curriculum Project Leader at the Assessment Curriculum and Technology Research Centre at the University of Melbourne.

Iddo Gal is Senior Lecturer and past-Chair, Department of Human Services Management, University of Haifa, Israel.

Ameena Ghaffar-Kucher is a Senior Lecturer and the Associate Director of the International Education Development Program at the University of Pennsylvania.

Suzanne Grant Lewis is the Director of the UNESCO International Institute for Educational Planning (IIEP).

Emily Hannum is a Professor of Sociology at the University of Pennsylvania, and specializes in education, child and youth welfare, and social inequality, particularly in China.

Rachel Hinton leads the Education Research Team at the UK Department for International Development.

Kathy Hirsh-Pasek is the Stanley and Debra Lefkowitz Faculty Fellow in the Department of Psychology at Temple University, Philadelphia, and is a Senior Fellow at the Brookings Institution, Washington DC.

Nancy H. Hornberger is Professor of Education at the University of Pennsylvania.

Sarah Howie is a Professor in the Faculty of Education, at the University of Pretoria.

Emmanuel (Manny) Jimenez is the Executive Director of 3ie, in New Delhi, and worked for 30 years at the World Bank Group.

Anil Kanjee is a Research Professor and Coordinator of the Postgraduate and Research Programme in the School of Education at the Tshwane University of Technology, Pretoria.

Marlaine E. Lockheed is a Visiting Lecturer in the Woodrow Wilson School at Princeton University, New Jersey, and previously worked at the World Bank.

Bryan Maddox is a Senior Lecturer in Education and International Development at the School of International Development, University of East Anglia, UK.

Silvia Montoya became the Director of the UNESCO Institute for Statistics (UIS) in 2015.

John Mugo is the Director of Data and Voice at Twaweza East Africa.

Joshua Muskin joined Geneva Global in November 2016 as Senior Program Director and Education Team Leader.

Moses Oketch is a Professor of International Education Policy and Development at the University College London Institute of Education.

Benjamin Piper is the Senior Director for Africa Education, at RTI International, based in Nairobi.

Lauren Pisani is Senior Specialist for Learning Research at Save the Children.

Abbie Raikes is an Assistant Professor and Director of Global Early Childhood Development at the University of Nebraska.

Pamela Robertson is a Research Fellow at the Assessment Curriculum and Technology Research Centre at the University of Melbourne.

Justin Sandefur is a Senior Fellow at the Center for Global Development.

Sylvia Schmelkes del Valle is Director of the Mexican National Institute of Educational Evaluation.

Jorge Sequeira was, until late 2016, Director of the UNESCO Regional Bureau for Education in Latin America and the Caribbean.

Dropti Sharma is Content Head (Early Literacy and Numeracy) at Pratham Education Foundation, New Delhi.

Amrit Thapa is Lecturer in the International Educational Development Program at the University of Pennsylvania Graduate School of Education.

Shaher Banu Vagh is a senior research fellow at the ASER Centre, New Delhi, and Pratham Education Foundation, Mumbai, India.

Dirk Van Damme is Head of the Innovation and Measuring Progress division at the Directorate for Education and Skills of the OECD, in Paris.

Daniel A. Wagner is the UNESCO Chair in Learning and Literacy, and Professor of Education at the University of Pennsylvania. He is the Director of the International Literacy Institute, and Penn's International Educational Development Program.

J. Douglas Willms is a Professor and Co-Director of the Canadian Research Institute for Social Policy at the University of New Brunswick, holds the Canada Research Chair in Literacy and Human Development, and is President of The Learning Bar.

Sharon Wolf is Assistant Professor in the Human Development and Quantitative Methods Division at the University of Pennsylvania Graduate School of Education.

Fei Yuan is Research Analyst in the Office of the Chief Economist, Africa Region, World Bank Group.

Learning at the bottom of the pyramid: an introduction

Daniel A. Wagner, Sharon Wolf, and *Robert F. Boruch*

Overview

The United Nations development goals have consistently placed a high priority on the quality of education – and on learning. This has led to substantial increases in attention to, and international development assistance for, the improvement of education worldwide. The development goals are mainly normative: that is, they tend to emphasize averages across nations, with relatively limited attention to variations within countries and to those performing at the low end of the distribution.

The *Conference on Learning at the Bottom of the Pyramid*, held in March 2017 at the University of Pennsylvania, aimed to focus attention on this issue. It brought together experts from around the world to explore the scientific tensions related to understanding learning among poor and marginalized populations in low-income countries – those at the 'bottom of the pyramid' or BoP (Wagner and Castillo, 2014).

International organizations, donor agencies, and many national governments often invoke populations at the BoP as the target of their investments – trying to help the poorest of the poor. Still, our understanding of learning – a key focus of the new 2030 UN Sustainable Development Goals (SDGs) – seems inadequate to the challenges ahead or ensuring learning for all. This concern is at the heart of the discussions that drew a group of learning and education specialists to the conference in 2017.

The origins of this effort date back to 1990, when the World Conference on Education for All in Jomtien, Thailand embraced two key educational goals: to significantly increase access to education for children in low- and middle-income countries (LMICs), and at the same time promote quality of education and learning. A decade later, at the 2000 Education for All (EFA) conference at Dakar, these same two challenges were expanded into a detailed list of six education targets in the Dakar EFA Framework for Action. The aims were to promote early childhood care, make primary school compulsory, address learning

needs for all, promote adult literacy, reduce gender disparities, and develop quality measures of learning outcomes (UNESCO, 2003: 28).

More recently, these international interests were reinforced in the UN Millennium Development Goals (MDGs) for 2015, where universal primary education was made the second of eight major goals (United Nations, 2000). These goals also led to greater awareness towards improving children's learning on a global scale. Thus, it was no surprise that in the next round of UN goals – the SDGs – learning was again at the centre of 17 broad goals to be achieved by 2030 (United Nations, 2015).

The growth in interest and support of children's learning has raised acute challenges, especially with respect to the BoP. Research on learning gaps among the poor in high-income countries has been a longstanding area of interest (e.g. Duncan and Murnane, 2014; OECD, 2012; Reardon and Portilla, 2016), and has been accompanied by substantial effort to create measurement tools for diverse populations (e.g. Dorans and Cook, 2016). Even so, the scientific community has, to date, invested only modest effort in understanding and narrowing learning differences in the BoP in LMICs. While some lessons can be learned from high-income countries, there are also unique issues in LMICs that warrant systematic research – issues that are highlighted throughout this volume.

The main purpose in this volume is to expand the conversation about learning for all in LMICs by bringing attention to marginalized communities within these countries. The premise is that focusing on average country-level performance (which is typically the focus of international agencies in achieving the UN goals) inevitably leaves out those who face the most barriers to education and learning. Thus, the aim here is to bring greater attention to reducing learning inequalities *within* countries, as a way not only to increase equity, but also to raise national levels of learning.

In this volume, a diverse group of authors discuss and analyse the scientific tensions in understanding learning among poor and marginalized populations in LMICs. Four broad areas are considered: how to define the BoP; how to measure and assess learning outcomes across diverse populations within a country; variations in learning across the life-span; and the implications for international education policy. Each of the 12 chapters is complemented by two commentaries, thus there are a total of 36 contributions.

Definitions (*Chapters 1–3*)

In the social sciences, *learning* is defined most commonly as a change – such as in knowledge, skills, attitudes, and values – based on experiences of some kind. Thus, schooling is not the same thing as learning. While schooling is usually designed to foster curriculum-based learning in classrooms, research increasingly demonstrates that much of what we presume is learned in school is not, and that a great deal of learning takes place outside of schools.

There is a large and diverse empirical research base in the area of human learning. However, much of the available research is limited by constraints of various kinds. Most prominent among these is the limited ability to generalize from findings in one population or context to others. In *Chapter 1*, Schmelkes considers common elements to human learning, in and out of school, including important cultural variations that are large and often poorly understood. She concludes that much more should be done to improve educational policy and address such contextual issues.

A second key priority is to determine what populations are meant by the phrase BoP. How do populations differ in LMICs – both across and within countries? As pointed out by Montoya in *Chapter 2*, there are at least six prominent dimensions through which populations at the BoP may be described in low-income countries, and each is important for considering the ways that young people can escape from persistent poverty. Building on the first two chapters, Crouch (*Chapter 3*) lays out a conceptual model, buttressed by data from international assessments, that describes how to flatten the learning pyramid to ensure more equitable learning outcomes for all by focusing on the poorest learners.

Overall, these three chapters provide a framework for considering the nature and extent of BoP studies of learning.

Measurement and assessment (*Chapters 4–7*)

There are many critiques of the educational assessment enterprise, the beginning of which is sometimes attributed to the French psychometrician Alfred Binet. In order to support the expansion of public schooling in France, Binet famously created assessments through which he could predict which children would have the most difficulty in school. In this section, we consider contemporary approaches to learning assessments, with a specific focus on the socio-cultural determinants of who succeeds and who does not at the BoP.

Kanjee, in *Chapter 4*, takes a broad perspective by reviewing the purposes of international assessment studies, suggesting that assessments have only limited impact on supporting BoP learning achievement. He concludes that assessments can better address the learning needs of poor and marginalized learners by reporting results through formative evaluations that can impact children before learning gaps widen. In *Chapter 5*, Willms describes a conceptual model for improved learning over the life-course, empirically supported by research in Uruguay among preschool children, and in Canada with young indigenous children. One of his findings is that in order to succeed in school, children need to learn to read with confidence during the primary grades, and use language to think critically, solve problems, and create new knowledge. He concludes that national and international assessments can serve to establish standards, assess the extent of inequalities among various sub-populations, and provide a framework for basic or theoretical research, but that there should be greater focus on changing classroom practice.

One way to understand the inner workings of assessments and use them to promote learning at the BoP is provided by Vagh and Sharma (*Chapter 6*) in their action research project in Allahabad, India. This project sought to develop and evaluate a local language literacy and numeracy programme for children from low socio-economic backgrounds in government school primary grades, using measurement and assessment to drive programme change. It highlights some of the strengths and challenges of localized assessments. Moreover, it suggests that assessments can be used by teachers to support early reading. Finally, in *Chapter 7*, Maddox asks a seminal question: To what extent are learning assessments able to identify and include individual and cultural differences, without reproducing relations of disadvantage? These issues are described in terms of test fairness and procedures for anticipating and removing sources of test bias. In a series of ethnographic studies, Maddox advises the reader to pay close attention to how assessments are carried out *in situ*, and how questions are interpreted by the person tested. Serious problems can and will ensue without such care in local contexts.

Age-related differences (*Chapters 8–10*)

Another approach to BoP issues is through a life-span perspective. How do measurement tools on learning and learning outcomes vary for young children, students in school, as well as among youth and adults? Three chapters in this section consider such age-related differences.

Dowd and Pisani, in *Chapter 8*, have been deeply involved in the field of assessments of young children before they reach school age. Their chapter reviews the application of the International Development and Early Learning Assessment (IDELA) instrument to explore young children's skills at the BoP and identify learning gaps in early academic, physical, and social-emotional development. Based on the broad findings from more than 20 LMICs, and closer analysis of particular contexts, the authors make that case that there is much variation in early childhood learning within countries, particularly between urban and rural contexts. They argue that national policies in support of early childhood need to be guided by disaggregated data in order to ensure that children at the BoP receive adequate support.

In *Chapter 9*, Care, Robertson, and Ferido describe how well-designed assessments for school-aged children can provide individualized information that can support school-based learning. These assessments build on the skill levels that children bring to the classroom. Through what they term a 'learning progression model', they present data on children in the Philippines who are best able to learn from specifically guided instruction tailored to their particular skill level. They conclude that learning assessments can and should be inclusive of diverse groups within any larger target population. Finally, Oketch (*Chapter 10*) focuses on youth and adult learning in sub-Saharan African, pointing out that rapidly changing demographics and economies in the region require significantly greater attention. Further, the population of low-skilled youth is growing dramatically, even though more African children are going to school than ever before. This chapter describes the importance of technical and vocational education and training (TVET) and non-formal education as two known methodologies for directly providing instruction and learning outside of the classroom in support of out-of-school youth and adults. The problem remains, according to Oketch, that there is a paucity of research in this domain, and in particular among populations at the BoP.

International policy perspectives (*Chapters 11 and 12*)

In *Chapters 1–10,* authors and commentators present multiple views on scientific definitions, measurement tools, and life-span approaches for understanding learning at the BoP. This final section of the volume considers the kinds of educational policy implications that need to be considered by both national and international decision-makers.

Benavot (*Chapter 11*) raises a key issue in supporting learning at the BoP, notably the need to move beyond easily accessible measures of learning – namely, school-based surveys of a narrow range of learning outcomes at the primary and lower secondary level – and engage with the broader and more comprehensive learning agenda proposed in the SDGs. He points out that many of the UN goals contain diverse elements of learning, and the specific targets for each goal may vary a great deal across diverse populations. Further, he notes that many of the key markers of disadvantage in education (such as socio-economic status, SES) are very difficult to change. He concludes that a serious focus on learning at the BoP will require greater clarity of definitions, and a more deliberate approach to building evidence on how best to improve relevant learning outcomes for the disadvantaged.

In *Chapter 12*, Van Damme provides a global policy perspective supported by the findings of the Organisation for Economic Co-operation and Development (OECD) international learning assessments. He asserts that international educational policies can only be inclusive and sustainable if those at the bottom of the social and educational pyramid benefit from them. To support this perspective, Van Damme presents findings that demonstrate how higher levels of economic growth are driven by more years of education and greater learning achievement within countries. By disaggregating data from the Programme for International Student Assessment (PISA) 2015 assessment, he reminds us that students with very low proficiency tend to drive down national averages (similar to the findings by Crouch in *Chapter 3*). He concludes that countries need to focus on raising average learning outcomes to desired national standards while at the same time narrowing the distribution of national learning outcomes.

Two key themes on learning at the BoP

The chapters and commentaries in this book cover broad conceptual issues, case examples, and differences across the life-span. Many ideas – some challenging – have surfaced. Two themes appear recurrent in and across these contributions that must be addressed for progress to be made in this area: sampling and comparability. We summarize how these themes intersect with the topics at hand.

Sampling. Population sampling is critical if we are to truly understand learning at the bottom of the pyramid. For example, about 5 per cent of the world's population resides in the United States, but nearly 95 per cent (Arnett, 2008) of scientific publications on

psychological development are based on populations that are WEIRD (Western, educated, industrialized, rich, and democratic; Heinrich, Heine, and Norenzayan, 2010) and living primarily in OECD countries. Moreover, of the research on psychological development conducted in the United States, about 80 per cent is on 'majority' ethnic groups (those of European origin), though these groups account for only about 50 per cent of the current US population (Arnett, 2008). More generally, much of the available research on learning is constrained in important ways by scientific data sets and research studies drawn from population samples living mainly within middle- to high-income countries. It seems obvious that researchers should explicitly address questions of representativeness and external validity, but often they do not. Fortunately, this state of affairs is beginning to change, as represented in this volume.

An additional sampling issue concerns that of language variation across ethnic groups, which exists in nearly every country, and especially in LMICs. Many of these groups, sometimes termed ethno-linguistic minorities, are well integrated into a national mix. For example, Latin America, with over 500 indigenous languages, is one region where intercultural bilingual education is expanding to promote social change. To date, 12 national governments have institutionalized multilingual pedagogy (Cortina, 2014). Often, social and political forces try to help resolve such differences, usually including policy decisions that result in a hierarchy of 'acceptable' languages to be used in schools and governance structures. In such situations, whether in OECD countries or LMICs, it is not unusual for children who are native speakers of minority languages to be excluded from research and assessments of learning. The 2010 Global Monitoring Report (UNESCO, 2010) describes how marginalization can seriously impede the educational attainment of these children. By excluding these children from learning assessments, their educational needs are less likely to come to forefront of a national education agenda.

Finally, consider the stakeholders who influence, oversee, or conduct the sampling of populations. Whether they are policy-makers, psychometricians, or local teachers, they all come to the task of sampling learning skills and populations with their own experiences and points of view. Choices about which skills to assess, among which populations, and in which languages, add potential bias to an already complex set of sampling issues. In order to address such biases, researchers can support the use of a range of methods including tailored sampling and subsample designs, matching samples, oversampling marginalized populations, and mixed methods designs.

Comparability. Comparability is central to global education databases, such as the large-scale data collection carried out by the UNESCO Institute for Statistics (UIS) and OECD (*see Chapters 2, 11 and 12*, this volume). Nonetheless, if the primary goal is comparability *across* countries, less attention may be paid to the assessment's validity of the definitions and classifications of learning *within* countries. In particular, such data may become less meaningful and potentially less applicable at the local level. Can both comparability and sensitivity to context be appropriately balanced in learning research? Should countries with below-average scores be tested on the same scales as countries that have much higher average scores?

If some countries, or groups of students, are located at the 'floor' of a scale, some would say that the solution is to drop the scale to a lower level of difficulty. Others might argue that the scale itself is flawed, and that there are different types of skills that could be better assessed, especially if the variations are evidently caused by cultural, ethnic, linguistic, and related variables that lead one to question the test as much as or more than the group that is tested. If the goal is to improve learning at the BoP, how credible (see *Chapters 3 and 7*, this volume) are the findings at the tail of the distribution from international (or even national) assessments?

The issue here is the degree to which it is necessary to have full comparability in learning outcomes, with all individuals and all groups on the same measurement scale. There is as yet no consensus as to whether localized and national monitoring assessments should also be internationally comparable. Yet, the merits and shortcomings of each need to be considered in discussions of learning for marginalized populations. If a choice is made not to 'force' the compromises needed for a single unified scale, what are the gains and losses in terms of comparability? Can international goals and commensurate statistics be maintained as stable and reliable if localized approaches are chosen over international comparability? See, for example, the responses to these questions in the work described in *Chapters 5, 6 and 9* in this volume.

In the end, international research on learning requires some form of comparability, but perhaps in more varied ways than are usually considered today. For example, international and regional assessments are aimed specifically at cross-national comparability, while hybrid assessments (*Chapter 6* of this volume; Wagner, 2011) are more focused on local contexts and increased validity. The latter try to combine aspects of large-scale and small-scale assessments, and may be thought

of as smaller, quicker, and cheaper. Which types of comparability are most important depends on the policy goals desired, as well as timing and resource considerations.

Conclusions

How to improve learning in LMICs and in poor and marginalized communities – bottom of the pyramid (BoP) populations – is the focus of this volume. What we mean by BoP, in terms of population characteristics, or in terms of learning, or in terms of policies to be addressed, was the focus of the 2017 Conference on Learning at the Bottom of the Pyramid.

It is reasonable to conclude that work in this area, especially in LMICs, has grown in scientific significance as well as for global education planners. Nonetheless, as we move forward from Jomtien, Dakar in 1990 to the UN SDG targets in 2030, it is clear that social and economic inequalities will persist unless a serious focus on learning among the poor is expanded. In his seminal book on new approaches for reaching BoP consumer markets, C.K. Prahalad (2006) challenged corporations to adopt a new philosophy of service delivery for this historically overlooked population. By transforming the way learning is understood in contexts at the BoP, we can begin to understand how to better promote educational quality and increase the learning consequences among those hardest to reach – what may be termed the 'learning equity agenda' (Wagner, 2018). Our hope is that this volume will provide new avenues to support this critical and much-needed global endeavour.

References

Arnett, J.J. 2008. 'The neglected 95%: why American psychology needs to become less American'. In: *American Psychologist, 63*(7), 602–614.

Cortina, R. (Ed.) 2014. *The education of indigenous citizens in Latin America*. Clevedon, UK: Multilingual Matters.

Dorans, N.J.; Cook, L.L. (Eds). 2016. *Fairness in educational assessment and measurement*. New York: Routledge.

Duncan, G.J.; Murnane, R.J. 2014. *Restoring opportunity: the crisis of inequality and the challenge for American education*. Boston, Mass.: Harvard Education Press.

Heinrich, J., Heine, S.; Norenzayan, A. 2010. 'The weirdest people in the world?' In: *Behavioral Brain Science, 33*(2–3), 85–135.

OECD. 2012. *Equity and quality in education: supporting disadvantaged students and schools.* Paris: OECD Publishing. Retrieved from: http://dx.doi.org/10.1787/9789264130852-en

Prahalad, C.K. 2006. *The fortune at the bottom of the pyramid: eradicating poverty through profits*. Upper Saddle River, N.J.: Wharton School Publishing.

Reardon, S.F.; Portilla, X.A. 2016. 'Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry'. In: *AERA Open*, 2(3). Retrieved from: http://ero.sagepub.com/content/2/3/2332858416657343

UNESCO. 2003. *Gender and Education for All – the leap to equality*. Education for All Global Monitoring Report. Paris: UNESCO.

——. 2010. *Reaching the marginalized*. Education for All Global Monitoring Report. Paris: UNESCO.

United Nations. 2000. United Nations Millennium Declaration. Resolution adopted by the General Assembly. United Nations A/RES/55/2. Retrieved from: www.un.org/millennium/declaration/ares552e.htm

——. 2015. Sustainable Development Goals. Retrieved from: www.un.org/sustainabledevelopment/sustainable-development-goals/

Wagner, D.A. 2011. *Smaller, quicker, cheaper: improving learning assessments in developing countries*. Paris and Washington: UNESCO IIEP and EFA Fast Track Initiative of Global Partnership for Education.

——. 2018. *Learning as development: rethinking international education in a changing world*. New York: Routledge.

Wagner, D.A.; Castillo, N.M. 2014. 'Learning at the bottom of the pyramid: constraints, comparability and policy in developing countries'. In: *Prospects*, *44*, 627–638.

Chapter 1
What is 'learning' in the case of marginalized populations in low-income countries?

Sylvia Schmelkes

Introduction

All children learn. Learning is synonymous with developing – if you are growing and developing, you cannot avoid learning (Ferreiro, 1998). And, as we now know, this phenomenon continues throughout life.

Children learn informally, most often through observation, imitation, and experimentation. They observe the context in which they live, and learn to give meaning to it (Rogoff, 2003). They imitate the adults around them, mainly their immediate family. With the tools at their disposal, they experiment with transformation: they learn to do. In rural areas, and among marginalized populations, this is the principal way in which children learn. These children have less access than others do to other sources of learning, such as media, the internet, books, theatres, and museums; they have fewer opportunities to make friends with children from other environments. All of these are also sources of informal learning; they open doors to other realities beyond the immediate ones.

Even though it is obvious, it is important to say that *any child that does not have an intellectual disability can learn anything,*[1] as Bruner (1996) informed us. However, external and school-based conditions impose limits on learning. In other words, there is an infinite potential for learning, but opportunities for doing so can be limited. Precisely because of this, it is a matter of concern that many children are not developing to their full potential. Understanding contextual limits to learning is, I believe, one of the important priorities of this book.

What is learned in school

Much has been written about what should be learned in school. The basic learning needs defined at the Jomtien meeting in 1990 are a fundamental

1. Children with intellectual disabilities also have an enormous potential for learning. In many cases, this affirmation is applicable to them also.

reference. The Jacques Delors report (1998) established four pillars of learning at the core of the purpose of education. More recently, UNESCO (2015) published *Rethinking Education*, which revisits both the Faure (1972) and the Delors reports from the perspective of a changing world, and from the need to re-humanize education. These pieces are crucial in the definition of what learning in school should be.

Formal schooling leads to learning that is difficult to achieve without a systematic methodology (pedagogy), which involves gradually increasing intellectual demands, eventually leading to the development of higher-order thinking skills. For this to occur, the mastery of languages is crucial. Two of the four operations of the mother tongue are achieved through informal learning: listening and speaking. Schooling only perfects them. However, the other two – reading with comprehension and writing for communication – can best be achieved through schooling, or otherwise require non-formal learning programmes to develop them. They are very difficult, though never impossible, to acquire through informal learning mechanisms.

Formalizing the four operations of mathematical language is also mainly achieved through systematized and graduated learning procedures that occur in school. The ability to solve context-based mathematical problems is achieved informally when access to schooling is denied (Ferreiro, 1983; Delprato and Fuenlabrada, 2008), or when teaching does not explicitly relate what is taught to context-related demands (Carraher and Carraher, 1985). However, the formalization of mathematical problem-solving through algorithms for adding, subtracting, multiplying, and dividing can best be achieved through schooling, or otherwise requires non-formal education programmes. Higher mathematics, such as algebra and calculus, are learned mainly in school, of course.

Schooling is also a privileged environment for learning to live together (Delors, 1998). School broadens the spectrum of diversity, allowing for close interaction with peers from different backgrounds, as well as with meaningful adults such as teachers and principals. Co-operative learning, teamwork for problem-solving, participatory exercises in democratic decision-making, non-violent conflict resolution, experimenting with civic responsibility over others, and developing the ability to teach others, are all necessary for integral personal and social development. All of these lead to value formation and are part of the socialization purposes of school.

All children should know their rights as human beings. They should be aware of the institutions in their country dedicated to the defence of human rights, as well as how to defend them for themselves and for others.

Identity, self-esteem, and a sense of belonging that goes beyond the family and immediate community have also traditionally been developed in school, though not always adequately or equitably. In the 19th century, nation-building was a very important function of national school systems that were being developed in many newly born countries, particularly in the ex-colonial developing world (Nassif, Rama, and Tedesco, 1984). Human beings need to understand their place in the world and in the cosmos, their moment in history, as well as their roots and their origins, of which they have to be made proud. School is perhaps the only place where this can be achieved.

School is also the place to democratize scientific knowledge and, perhaps more importantly, the scientific way of knowing. Discovering the place one occupies in the world and in the universe is one of the basic learning needs. Understanding how our body works and how to take care of it and prevent infirmities and epidemics is also basic. The 2030 agenda has brought to light the importance of education for sustainable development as a crucial contribution of school to ensuring human life persists on the planet.

Moreover, schooling is the process through which children should learn how to learn, and learn to enjoy learning. More than learning specific content, which continuously increases by leaps and bounds and rapidly becomes obsolete, it is necessary to be able to learn to access information and discriminate it, and learn to access knowledge, as well as to distinguish when expert knowledge is necessary or when what we discover is sufficient for our purposes. In the process of learning to learn, children must discover that learning is enjoyable, and thus develop the motivation to continue learning throughout life.

Finally, school is the place where higher-order thinking is developed. Piaget explained how concrete thinking naturally develops in humans, even without schooling (1972). However, with exceptions, only systematic and gradual access to learning challenges leads to the development of abstract thinking. Kohlberg (1981) demonstrated how this is also true for moral development. With exceptions, lack of quality schooling explains why many people do not go beyond stages 2 (the Talion law) or 3 (the Golden Rule) of the six that he describes. Higher-order moral development, however, is necessary for responsible democratic participation and civic

action, and key to fostering respect for others and, ultimately, peaceful coexistence.

All children around the world should at least have access to everything that I have listed so far. This is why schooling is a basic human right. These are the basics for all children, and should be no different for marginalized children in the developing world.

Contextualized, relevant learning

In addition to what has been listed above regarding what should be learned in school, children should be able to relate what they learn to their context, to find a deep understanding of their immediate world along with the tools for its care and transformation, and discover their own culture in their native language. There should be a strong relationship between what children learn in school and what is needed to make learning relevant and meaningful to the children of the very diverse cultures that populate the world.

As an example, a recent project carried out among indigenous peoples in Mexico (Instituto Nacional para la Evaluación de la Educación, 2016*a*) informs us that indigenous communities highly value what they are taught in school – they clearly realize the necessity of relating to and getting along in the outside world. However, they feel that what they learn in school does not relate to or respect their community's norms and values. They complain about the fact that their language is not used in school, and is even forbidden in some cases. Teachers in general ignore the knowledge of the indigenous peoples, their customs and traditions, and their productive practices. Schooling targets individual learning, whereas indigenous peoples value communitarian attitudes and practices. Unwittingly perhaps, schooling causes indigenous students to develop a lack of interest in their own culture, and diminishes their desire to participate actively in it. When asked about the characteristics they would like the schooling to instil, they describe a value of their own culture and a willingness to participate in it, as well as preparation to face – and develop in – external contexts, and the ability to establish horizontal dialogue with members of the dominant and other cultures. Schools, in addition to ensuring that their students learn the basics – Spanish, literacy, and basic mathematics – should respond to the learning needs of the community. Students should also achieve mastery of their mother tongue. School should promote the values of the community – respect for elders, for others, and for nature is specifically mentioned – incorporate community knowledge and their

own ways of learning, and use human and natural resources from the community as educational inputs. Indigenous communities would like to have a greater say regarding what goes on in school, and would like to play a role in supervision, teaching, and assessment.

For Tomasevsky (2003), the right to education involves four A's: availability, accessibility, acceptability, and adaptability. Whereas the first two relate to the right to education, the second two relate to children's rights within education. Acceptability must be judged from the perspective of the learner – school has to be a place where students feel respected and secure. They must also feel that what is learned is both useful and meaningful, both for the present and for the future; both for living within the community and for relating to the outside world.

What is actually learned in school

The paradox is that basic education should ideally be highly demanding, however we know that many children are actually learning very little in school, especially those living in developing countries and in marginal areas. This we can only judge by the results of standardized tests,[2] both national and international, and mostly we can only speak of reading, arithmetic, and, in some cases, science. According to these tests, we know that children in these environments are not learning these basics. In TERCE 2013 (UNESCO, 2016*a*), a Latin American comparative assessment, at the end of primary school, one tenth of the children in Mexico (above the mean of the participating countries) and 15 per cent in Guatemala (below the mean) were unable to understand a narrative text. Some 23 per cent of Mexican children and 33 per cent of Guatemalan children were unable to operate with the four basic algorithms (see *Figure 1.1*).

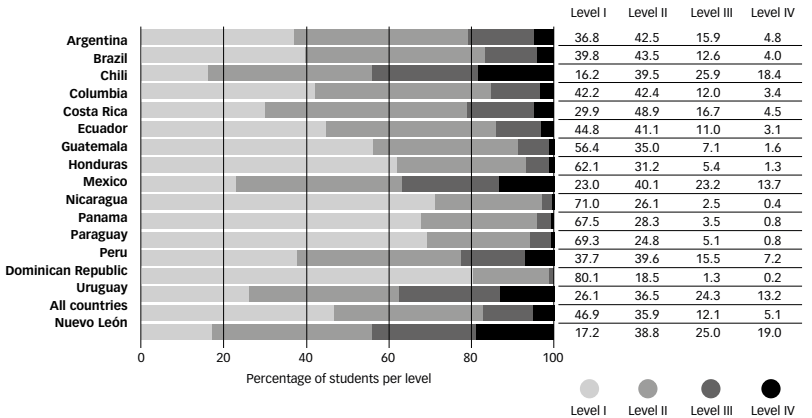
It is interesting to note that the factors associated with these results belong to both the demand and the supply side (UNESCO, 2016*b*). The socio-economic status (SES) of the family is an important factor, but so is having repeated a grade, for example.

National tests in Mexico reveal more alarming results (INEE, 2016*c*). At the end of lower secondary school (grade 9), 65 per cent of

2. There is a strong debate about whether standardized tests are the most adequate instruments to measure learning among children living in disadvantaged conditions, and more so when teaching is carried out in a language different from the one spoken by the child. We can even question whether poor results are not at least partly due to the characteristics of these tests when applied to students in these circumstances. See further discussion in this volume by Kanjee.

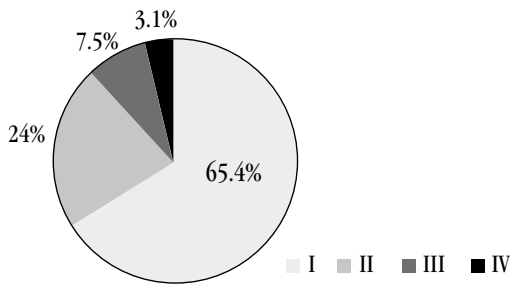
Mexican children can only solve problems that involve using the four basic operations with integral numbers. This is true of 84 per cent of the children that attend community secondary schools, that is, those that are located in the more remote and dispersed rural areas (see *Figure 1.2*).

Figure 1.1 Distribution of Grade 6 students, according to achievement in mathematics



Source: UNESCO, 2016*a*.

Figure 1.2. Percentage of Grade 9 Mexican students, according to level of achievement



Source: INEE, 2017.
Notes on levels: N I. Students show insufficient learning and will probably have severe difficulties in continuing their studies. N II. Students show very basic learning, the minimum necessary to continue their studies. N III. Students have achieved adequate learning. N IV. Students excel in their achievement.

PISA is a more demanding test because it measures competencies that involve higher-order skills, and requires students to apply what they know to new situations that relate to the demands of global society.

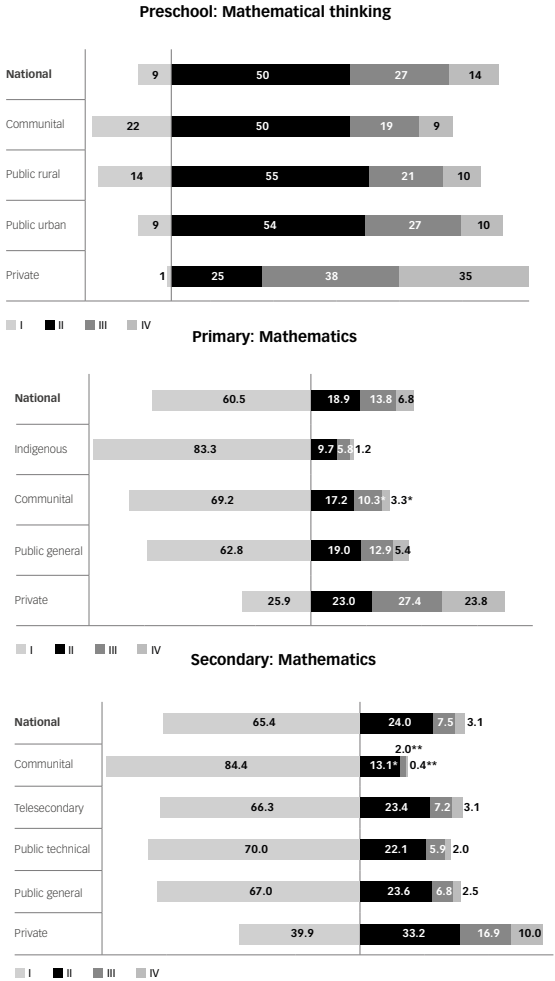
OECD first administered PISA only in OECD countries, but gradually expanded to other countries around the world, including a few developing countries. PISA scores are arranged in five or six levels, depending on the test. Level 2 is considered the basic minimum needed to be able to face the demands of modern society. Students scoring at level 1 are below this basic minimum. However, the percentage of students in developing countries scoring at level 1 was so large that it had to be divided in two (1a and 1b) as of 2006, in order to be able to say something meaningful about what children at level 1b actually know and are able to do. Even after doing that, in PISA 2015, 1 per cent of Mexican students and 16 per cent of Dominican students remained below level 1b (INEE, 2016*b*). Almost half of Mexican students achieved below level 2. This is true of 86 per cent of Dominican students.

Why are we not seeing learning take place in marginal schools in developing countries?

The most interesting result of all standardized tests is the fact that there are always small percentages of students and schools that achieve at levels 3 and 4 of the different scales involved in different tests. This is true despite the level of marginality of the locality, or how low the SES of the family; no matter whether the students belong to a minority culture or speak a different language than the dominant one in the country. This demonstrates the thesis that we started out with: any child can learn. It also shows that schools can make a difference. This is precisely why *standards must remain high*.

It is important to find out what explains these exceptions to the rule that the poorer you are, the less you learn in school. The effective school movement has shown us some of the things that are common in schools that have exceptional achievement rates in disadvantaged conditions. Meta-studies of effective schools (Sammons, Hillman, and Mortimore, 1995) find that these schools have good instructional leadership, have developed good team-work among teachers, create an adequate climate where children feel accepted and respected, have internalized a culture of planning and of monitoring goal-attainment, have high expectations of their students, and relate well with parents and the community. These meta-studies also show that each school is different. Although the characteristics just listed should be in place in order to expect good student achievement results, there are many other variables that differ among schools, which respond to both context issues and teacher orientation and talent.

Figure 1.3 Mexico: learning results at the end of the three compulsory levels, according to types of schools

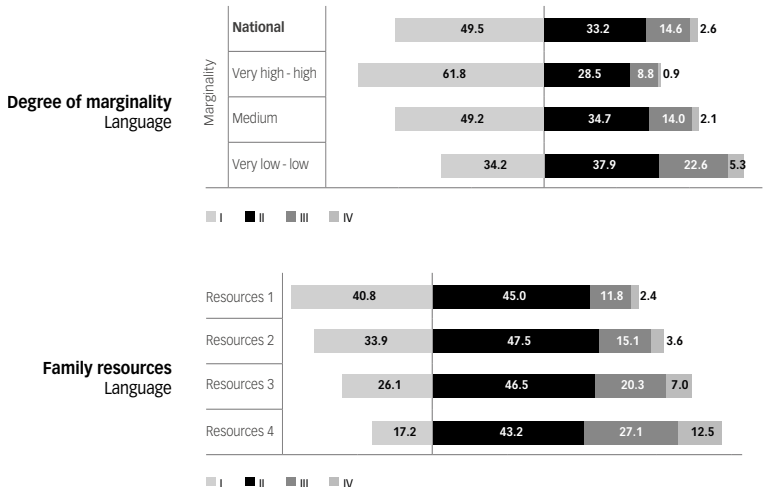


Source: INEE, 2016*c*.

We have indicated several other requirements for good schools. A very important one is for the school and the teacher to innovate and adapt teaching to context and cultural characteristics in order to make learning meaningful. I have also pointed out the importance of the language of instruction and communication, which should be that of the child's mother tongue when developing mastery, as well as for the adequate acquisition of a second language. Schools have to be accessible in the sense that no physical, economic, or socio-psychological barriers (especially discrimination)

should be in place (Tomasevsky's second 'A'). Obviously, schools should have adequate infrastructure and be sufficiently equipped in such a way that teaching takes place in acceptable conditions and that learning can flourish. We know, however, that this is not the case of schools located in disadvantaged areas, where unfortunately poverty on the demand side meets with poverty on the supply side, and poor results are probably in a large part explained by this phenomenon (see, for example, INEE, 2016d)[3].

Figure 1.4 Mexico: learning results in Grade 6, according to marginality of locality and family resources



Source: INEE, 2016*c*.

I have said little about an adequate pedagogy. It is hard to do so when we have stated that ways of learning differ between cultures and that teaching should adapt to these different ways of learning. Nevertheless, some pedagogical principles seem to be applicable universally (Borphy, 1988). Students should become agents of their own learning, which means that emphasizing activities that involve discovering, observing, is perhaps, together with learning to live together, the most important purpose of

3. The schooling model that is present worldwide emerged for urban middle-class children. The same model has gradually been taken elsewhere, to rural areas, cultural minorities, and marginalized populations. In so doing, it has generally suffered: while the urban model assumes that there is one teacher for each group, in the rural areas multigrade schools are prevalent; the urban model requires a full-time manager in the form of a director or headmaster, while in many small schools one of the teachers takes on this role as well as teaching.

formal schooling. Children learn together, and often children learn more from other children who are perhaps one or two steps ahead of them in cognitive development than from the teacher. Thus, students should be encouraged to teach other children who have remained behind, as this kind of peer-to-peer instruction also fosters learning in those who teach. Mistakes should be considered opportunities for improving learning. Learning involves asking questions more than learning answers. Learning, and particularly socio-affective learning, takes place through dialogue and discussion, for which opportunities in the classroom ought to be frequent. Metacognition activities lead to the discovery of how one learns, and make transference to other learning demands possible. Children should be continuously challenged, so that they find new material exciting and new activities enjoyable. Listening to and taking into account what children have to say about whether they enjoy what they are learning and the way they are supposed to do so has been shown to improve teaching. All of these are of course open to debate, and many more principles can surely be added to the list. But the point is that poor pedagogy is undoubtedly one of the most important reasons for poor learning. We can hypothesize that the poorer the environment, the less these pedagogical principles are operational in the classroom.

Concluding remarks

In finalizing, I would like to summarize the points that I am raising for discussion regarding the learning of marginalized children in developing countries:

1. Any child – excepting those with severe intellectual disabilities – can learn anything. Children have to learn many things in school. These are constituent of their right to education. This is why we cannot lower expected standards – on the contrary, we must keep them high.
2. An important percentage of children in disadvantaged environments and in developing countries are not learning what standardized tests measure because of external conditions. External conditions limit learning.
3. External limitations to learning emerge from two sources: demand-side factors and supply-side factors.
4. The factors from the demand side have to do with the wider environment. Poverty is perhaps the most important; it often leads to child labour, which is another crucial factor, as well as to poor nutrition and vulnerability to sickness that keeps children out of school. Distance from the school (which affects mainly girls and

very young children) is another. Parental expectations regarding schooling are a third factor. The cultural distance of the family and community from the culture of school is also crucial.

5. The factors that come from the supply side have to do, among other things, with the physical conditions of the school, the culture of the school (regarding planning, monitoring, climate, and the relationship to the parents and the community), the use of the mother tongue as both an object of learning and the language of instruction and communication, and the knowledge and the pedagogy of the teacher.
6. The most important factor from the supply side, given basic teaching and learning conditions, is the ability to make teaching relevant and learning meaningful. This involves the participation of the immediate community, the ability of the school director to develop an adequate school culture, and the adequate training and pedagogical performance of teachers. Some conditions demand intersectoral policies that are able to mitigate poverty and to combat child labour, as well as tend to malnutrition and health issues.

We are seeing poor learning results on standardized tests from children living in marginal areas and in developing countries because we are looking only at the result, when contextual conditions are problematic and pedagogical processes are flawed. Not nearly enough is being done, on the policy side, to improve the contextual conditions and the procedural factors that lead to poor learning.

Much of the above is still hypothetical. We have yet to demonstrate in most cases that modifying these factors leads to improved learning outcomes. However, we cannot continue to rely solely on the measurement of learning outcomes as the way to judge the world's progress in improving the quality of education. We have to also measure whether we are improving the factors that we believe determine poor results and prevent the observance of the right to quality education of many children around the world.

References

Brophy, J. 1998. *Teaching*. Geneva: International Academy of Education, International Bureau of Education and UNESCO.

Bruner, J. 1996. *The culture of education*. Cambridge, Mass. and London: Harvard University Press.

Carraher, T.N; Carraher, D.W. 1985. *En la vida diez, en la escuela cero*. Mexico City: Siglo XXI.

Delors, J. 1998. *Learning, the treasure within. Report to UNESCO of the International Commission on Education for the Twenty-first Century*. Paris: UNESCO.

Delprato, M.F.; Fuenlabrada, I. 2008. 'Así le hacemos nosotras: Prácticas de numeración escrita en organizaciones productivas de mujeres con baja escolaridad'. In: *Cuadernos de Educación, 6*(6).

Faure, E. 1972. *Learning to be: the world of education today and tomorrow*. Paris: UNESCO.

Ferreiro, E. 1983. *'Los adultos no alfabetizados y sus conceptualizaciones del sistema de escritura'*. Mexico City: Departamento de Investigaciones Educativas, Centro de Investigación y Estudios Avanzados.

——. 1998. *Caperucita roja aprende a escribir: estudios lingüísticos comparativos en tres lenguas.* Mexico City: Gedisa.

INEE (Instituto Nacional para la Evaluación de la Educación). 2016*a*. *Consulta previa, libre e informada a pueblos y comunidades indígenas sobre evaluación educativa. Informe General de Resultados.* Retrieved from: http://publicaciones.inee.edu.mx/buscadorPub/P1/D/245/P1D245.pdf

——. 2016*b*. *México en PISA 2015*. Mexico City: INEE.

——. 2016*c*. *La Educación Obligatoria en México: Informe 2016.* Mexico City: INEE.

——. 2016*d*. 'Infraestructura, mobiliario y materiales de apoyo educativo en las escuelas primarias. ECEA 2014'. Retrieved from: http://publicaciones.inee. edu.mx/buscadorPub/P1/D/244/P1D244.pdf

——. 2017. *El aprendizaje de los alumnos de sexto de primaria y tercero de secundaria en México: Lenguaje y comunicación y Matemáticas, Informe de resultados Planea 2015.* Mexico City: INEE.

Kohlberg, L. 1981. *The meaning and measurement of moral development*. Atlanta, Ga.: Clark University Press.

Nassif, R.; Rama, G.; Tedesco, J.C. 1984. *El sistema educativo en América Latina.* Santiago de Chile: UNESCO, Economic Commission for Latin America (CEPAL), and UN Development Programme.

Piaget, J. 1972, 'Intellectual evolution from adolescence to adulthood'. In: *Human Development, 15*(1), 1–12.

Rogoff, B. 2003. *The cultural nature of human development*. Oxford: Oxford University Press.

Sammons, P.; Hillman, J.; Mortimore, P. 1995. *Key characteristics of effective schools: a review of school effectiveness research*. London: Ofsted.

Tomasevski, K. 2003. *Education denied*. London: Zed Books.

UNESCO, 1990. *Meeting basic learning needs*, report on World Conference on Education for All. Paris: UNESCO.

——. 2015. *Rethinking education: towards a global common good?* Paris: UNESCO.

——. 2016*a*. *Tercer Estudio Regional Comparativo y Explicativo. Informe de Resultados.* Laboratorio Latinoamericano de Calidad Educativa. Santiago: UNESCO Regional Education Office for Latin America and the Caribbean. Retrieved from: http://unesdoc.unesco.org/images/0024/002435/243532S.pdf

——. 2016*b*. *Tercer estudio regional comparativo y explicativo. Factores Asociados.* Santiago: UNESCO Regional Education Office for Latin America and the Caribbean. Retrieved from: http://unesdoc.unesco.org/images/0024/002435/243533s.pdf

Commentary

Benjamin Piper

Achievement gaps in developing countries are an under-researched phenomenon. In the United States, on the other hand, funding for research and interventions has been targeted at identifying the contours of the country's gap (i.e. ethnicity, wealth, and location). Moreover, in pre-service courses and practica, colleges of education routinely train teachers to consider methods that narrow the gap. The result has been a modest decline in the magnitude of racial achievement gaps in US results on the National Assessment of Educational Progress (NAEP) (McFarland *et al.,* 2017). That such costly interventions have resulted in only a limited contraction of the gaps is disconcerting, given the widening achievement gaps across ethnicity, wealth, and location identified in Latin America via the 2006 Second Regional and Comparative Explanatory Study (Segundo Estudio Regional Comparativo y Explicativo, or SERCE), as well as those in sub-Saharan Africa revealed by the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ).

Schmelkes puts the onus of change on the system, and emphasizes both improved pedagogical methods and a broader understanding of learning. Given her perspective, I wonder whether Schmelkes would laud the modest reduction in the US achievement gap. The NAEP evidence suggests slight improvements in specific basic skills, rather than better performance on a more expansive set of measures estimating the broader cultural skills discussed in her essay. In developing-country contexts, there is a tension between the desire to have contextualized knowledge specific to each particular minority community, on the one hand, and the core difficulty of offering education of even moderate quality to poor and minority populations at scale, on the other. Given current funding levels, can countries do both?

'Poor pedagogy is undoubtedly one of the most important reasons for poor learning', argues Schmelkes. Supporting teachers to provide equitable educational opportunities in poor and wealthy schools in the basic areas of literacy and numeracy might have an opportunity cost that obviates efforts to expand the cultural relevance of the curriculum. There are encouraging examples of high-quality, culturally responsive teaching (Colbert and Arboleda, 2016), but do such programmes also equitably improve access to basic skills *at scale*?

Recent advances in assessment tools for those at the bottom of the pyramid have allowed us to identify the extent of the achievement gap in developing countries, and these gaps yawn wide in rural sub-Saharan Africa and Asia. Initial evidence from large-scale interventions focused on the gap has suggested that while the poor and minorities have better outcomes as a result of the interventions, persistent interaction effects between treatment and poverty mean that the non-poor actually *benefit more* from programmes designed to help the poor (Piper, Jepkemei, and Kibukho, 2015; Riecken and Boruch, 1974). Stated another way, there is a perverse and persistent 'Matthew effect' (where the rich get richer, and the poor lose even what they have) (Stanovich, 1986).

Schmelkes points to a time when it will be possible for countries to teach basic skills *and* ensure cultural relevance. Where governments currently struggle to provide either, how do policy-makers focused on improving outcomes for millions, not just thousands, choose between more basic skills, more contextualized learning, or just a little of both? These are impossible and unfair choices, surely, and Schmelkes' paper prods us to consider them urgently and consider the perspectives and desires of the community in these fraught decisions.

References

Colbert, V.; Arboleda, J. 2016. 'Bringing a student-centered participatory pedagogy to scale in Colombia'. In: *Journal of Educational Change,* published online 5 November 2016. https://doi.org/10.1007/s10833-016-9283-7

McFarland, J.; Hussar, B.; de Brey, C.; Snyder, T.; Wang, X.; Wilkinson-Flicker, S.; Gebrekristos, S.; Zhang, J.; Rathbun, A.; Barmer, A.; Bullock Mann, F.; Hinz, S. 2017. *The condition of education 2017.* Report no. NCES 2017-144. Washington DC: Institute for Education Sciences, National Center for Education Statistics. Retrieved from: https:// nces.ed.gov/pubs2017/2017144.pdf

Piper, B.; Jepkemei, E.; Kibukho, K. 2015. 'Pro-poor PRIMR: improving early literacy skills for children from low-income families in Kenya'. In: *Africa Education Review, 12*(1), 68–85. Retrieved from: https://doi.org/10.1080/ 18146627.2015.1036566

Riecken, H.W.; Boruch, R.F. 1974. 'Experimental assessment of special educational programs'. In: H.W. Riecken and R.F. Boruch (Eds), *Social experimentation: a method for planning and evaluating social intervention* (pp. 301–308). New York: Academic Press.

Stanovich, K.E. 1986. 'Matthew effects in reading: some consequences of individual differences in the acquisition of literacy'. In: *Reading Research Quarterly, 21,* 360–406.

Commentary

Sharon Wolf

As Schmelkes notes in her paper, 'All children learn'. This is a critical premise from which to begin the conversation on 'learning at the bottom of the pyramid'. In academic scholarship, learning is defined as *a modification of behaviour due to experience*, and may reflect changes in knowledge, behaviours, skills, attitudes, or values. Learning is essential to reap the benefits of education. As Schmelkes makes clear, the *learning process* is not different among marginalized populations. What is different are the opportunities for learning.

A needed emphasis on learning across multiple domains

There are multiple and interrelated domains of learning. Schmelkes notes that school develops a broad range of skills, including academic, social, and personal. At different stages of development, different skills become more or less central to the learning process. Non-academic skills, such as social-emotional and executive function skills, facilitate children's learning (Durlak *et al.*, 2011). When children can regulate their emotions and behaviours and sustain their attention, they benefit more from learning opportunities (Raver, 2002). For youth, a sense of identity, life skills, and self-esteem support learning to ensure a healthy transition into adulthood.

Yet in low-income countries (LICs), and among marginalized populations in particular, education and learning are often measured by access to schooling, attendance, gender parity, or, in more recent years, by student achievement test scores. The multiple domains of learning must be considered if learning is to be understood and promoted among marginalized populations.

Learning inside and outside of school

Very little is known about how individual families prepare their children for school in LICs. Indeed, nearly all of the research on the relationships between household characteristics like socio-economic status (SES), parental investments, and child development has taken place in North America and Western Europe (Bornstein *et al.*, 2012). Better understanding of the nature of interactions between children and their caregivers is needed to promote learning among marginalized groups. The development of the Indigenous Motivational Caregiving

Practices model as suggested by Wadende, Oburu, and Morara (2016) provides an example in East Africa. This model suggests that caregivers engage in many unwritten practices based on culturally specific expectations of children, such as cleaning themselves or babysitting siblings (Nsamenang, 2008). These skills build children's sense of self-efficacy, and are highly scaffolded. Such a framework of informal learning, as Schmelkes describes it, should be incorporated to build on the investments parents are already making in their children's learning.

Ensuring that maintaining high standards does not exacerbate inequality

Maintaining high standards for all schools and all children is critical. However, demanding such standards without providing schools and teachers with the proper supports may result in unintended consequences. In the United States, for example, more stringent standards have not been met with increased school quality or increased learning outcomes, as schools serving marginalized populations are insufficiently resourced (Duncan and Murnane, 2014). Well-resourced schools, on the other hand, can ensure their students meet the high standards. Inadvertently, this policy exacerbated the consequences of differences in schooling quality and the effects of income inequality on the life chances of low-income children.

References

Behrman, J. R.; Schott, W.; Mani, S.; Crookston, B.T.; Dearden, K.; Duc, L.T.; Fernald, L.C.H.; Stein, A.D. 2017. 'Intergenerational transmission of poverty and inequality: parental resources and schooling attainment and children's human capital in Ethiopia, India, Peru, and Vietnam'. In: *Economic Development and Cultural Change, 65*(4), 657–697.

Bornstein, M.H.; Putnick, D.L. 2012. 'Cognitive and socioemotional caregiving in developing countries'. In: *Child Development, 83,* 46–61.

Duncan, G.J.; Murnane, R.J. 2014. 'Challenges in the classroom'. In: G.J Duncan and R.J. Murnane, *Restoring opportunity: the crisis of inequality and the challenge for American education* (ch. 4). New York: Russell Sage Foundation.

Durlak, J.A.; Weissberg, R.P.; Dymnicki, A.B.; Taylor, R.D.; Schellinger, K.B. 2011. 'The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions'. In: *Child Development, 82*(1), 405–432.

Fernald, L.; Schnaas, L.; Neufeld, L.; Knauer, H.; Guerra, A.G. 2014. 'Adding a parental support intervention to conditional cash transfers improves child development in rural Mexico'. In: *FASEB Journal, 28*(1), 378–8.

Nsamenang, A.B. 2008. '(Mis)understanding ECD in Africa: the force of local and global motives'. In: A. Garcia, A. Pence, and J.L. Evans (Eds), *Africa's future, Africa's challenge: early childhood care and development in sub-Saharan Africa* (pp. 135–141). Washington DC: World Bank.

Raver, C.C. 2002. 'Emotions matter: making the case for the role of young children's emotional development for early school readiness'. In: *Social Policy Report, 16*(3), 3–24.

Wadende, P.; Oburu, P.O.; Morara, A. 2016. 'African indigenous caregiving practices: stimulating early childhood development and education in Kenya'. In: *South African Journal of Childhood Education, 6*(2), 1–7.

Chapter 2
What is the bottom of the pyramid in the case of low-income countries?

Silvia Montoya

Introduction

Aid to reduce poverty has increased over the last two decades.[4] But a high percentage of the global population continues to live in poverty, especially in low-income countries (LICs).

In 1997, the General Assembly of the United Nations gave special attention to poverty, proclaiming the First United Nations Decade for the Eradication of Poverty (1997–2006). The 2000 Millennium Declaration sought to combine the numerous efforts to address poverty and economic and social development in a holistic fashion. Two years later, it was complemented by the International Conference on Financing for Development (2002), which was approved through the Monterrey Consensus. In 2015, countries adopted the Sustainable Development Goals (SDGs) to end poverty, protect the planet, and ensure prosperity for all.

Organizations and donors traditionally use gross domestic product (GDP) per capita to guide aid and identify LICs. But what would happen if we took additional variables into account when calculating this measurement, such as poverty, access to education, health, or the context in which people live? This chapter seeks to identify profiles and characteristics of countries in vulnerable situations measured through a set of socio-economic variables.

Who comprises the base of the pyramid?: exploring the concept

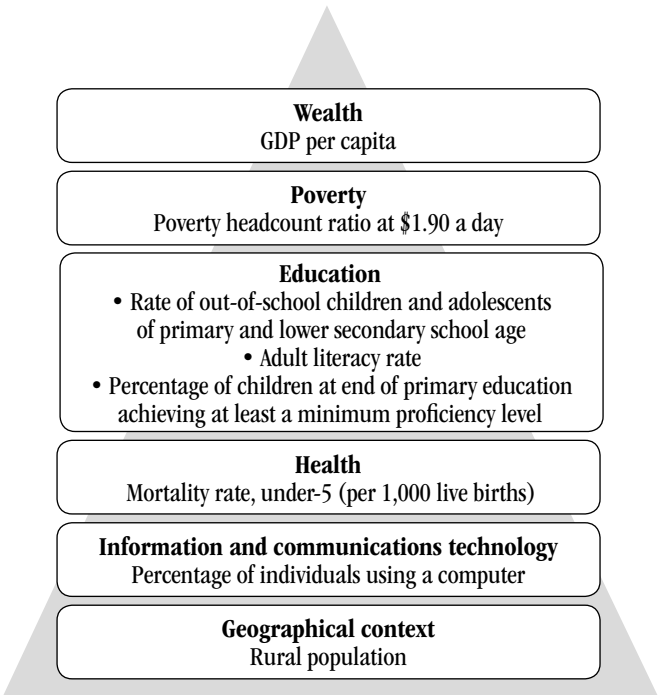
For statistical purposes, the methodology proposed by the World Bank currently divides national economies into four groups based on their gross national income (GNI) per capita (in US dollars): low, lower middle, upper middle, and high. National development agencies and the Development Assistance Committee (DAC) of the OECD use this type of classification

4. I wish to thank Jhonn Espinoza and Robert Boruch for their comments on a previous version of this paper.

to allocate aid from donor countries and assess the eligibility of each beneficiary country. This type of measurement has been used since 1970. Currently, Official Development Assistance (ODA) focuses on LICs or territories. But middle-income groups also receive help.

The income structure of the world population, or of a country, has the shape of a pyramid. Low-income countries, or the poor sectors within a country, form the base of the pyramid. As one ascends the pyramid, the number of countries or a population decreases, while individual income increases. In the past, when a country's income pyramid had a large proportion of the population living below the poverty line, investment expectations have been low. Yet, since the publication of *The Fortune at the Bottom of the Pyramid* (Prahalad, 2005), countries with large poor populations have increasingly been seen as offering an investment opportunity, where rapid growth and poverty reduction could be achieved with profitability, as many private, large-scale projects aimed at the BoP have proved profitable. The characteristics

Figure 2.1 Measurement dimensions of the BoP



Source: Prepared by the author.

of these projects have been that the products or services offerred were: (1) oriented towards the population at the BoP; (2) adapted to a geographic context; (3) priced so as to be accessible to the population at the BoP; and (4) innovative.[5]

While it is possible to use only GDP per capita to identify the countries at the BoP, countries have other socio-economic and geographic dimensions to take into account for the success of aid and investment. Under this premise we propose to analyse six dimensions described in *Figure 2.1*.

Who comprises the base of the pyramid?: the dimensions

Wealth dimension

The GDP per capita indicator is available for 182 countries. The maximum value is $141,543 (Qatar) and the minimum is $619 (Central African Republic). Using this indicator, the 46 countries that are located in the poorest/lowest quartile in our analysis could represent the BoP.

Although this income classification shows the average differences between countries, this indicator does not identify the distribution of the population according to their income or the concentration of wealth or poverty.

Poverty dimension

A few decades ago, inequality was not considered a pressing concern for the international agenda, since the focus was mainly on economic growth. But in recent years, the inequality gap both across and within countries has expanded exponentially. Low-income countries have a wide gap in their pyramid structure. In 1920, the income ratio between the richest 20 per cent of the planet's population and the poorest 20 per cent was 7 to 1; in 1960 it was 30 to 1; in 1980, 60 to 1; and in 2000 it rose to 80 to 1 (Dubois, 2009).

5. The population at the BoP showed a predisposition for learning and change. These qualities can be verified in the success of micro-credit projects for low-income families in Colombia, the use of fertilizers to improve the production of small farmers in Rwanda, the use of solar panels aimed at low-income households in Nicaragua, or the business network of small farmers in India connected through personal computers to discuss prices and decide for how much and when to sell their products. On the one hand, the user or consumer from the BoP had some basic knowledge to identify the benefits of the services or products purchased. On the other hand, the innovation of these enterprises required small local partners and employees from the BoP to sustain and expand commercial and production activities on a large scale. In general, the application of innovations and new technologies requires a minimum of basic knowledge from the users. In this context, BoP education plays an important role. A greater number of years of schooling at the BoP results in savings of time and resources allocated to the training of clients and employees.

Table 2.1 GDP per capita by purchasing power parity (PPP) by country with available information (current international $), lowest quartile, 2015

Country or territory	Indicator	Country or territory	Indicator
Central African Republic	619	Mali	2,028
Burundi	727	Benin	2,057
DR Congo	784	Chad	2,176
Liberia	835	Solomon Islands	2,201
Niger	955	Timor-Leste	2,399
Malawi	1,184	Senegal	2,421
Mozambique	1,192	Nepal	2,462
Guinea	1,209	Tanzania	2,673
Guinea-Bissau	1,456	Yemen	2,821
Togo	1,460	Tajikistan	2,834
Madagascar	1,465	Lesotho	2,950
Comoros	1,483	Vanuatu	2,988
Sierra Leone	1,569	Kenya	3,089
Ethiopia	1,629	Cameroon	3,115
Gambia	1,680	Sao Tome and Principe	3,219
Burkina Faso	1,696	Bangladesh	3,340
Haiti	1,757	Kyrgyz Republic	3,434
Rwanda	1,762	Cambodia	3,490
Zimbabwe	1,787	Djibouti	3,491
Uganda	1,851	Micronesia, Fed. Sts.	3,497
South Sudan	1,854	Côte d'Ivoire	3,514
Afghanistan	1,925	Zambia	3,836
Kiribati	1,995	Marshall Islands	3,911

Source: Prepared by the author based on the data published on https://data.worldbank.org/, accessed 14 April 2017.

The information needed to measure the concentration of wealth or poverty of the population is very scarce. In the poverty dimension, we propose the poverty headcount ratio at $1.90 a day. This variable shows the percentage of the population living on less than $1.90 per day. It is important to note that there is only data available for 89 countries or territories. The maximum value is 77.8 per cent (Madagascar) and the minimum is 0 per cent (Moldova). If we select only countries in the highest quartile, 23 countries would be located in the BoP (*see Table 2.2*).

Table 2.2 Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population) by country, highest quartile, latest available year between 2009 and 2014

Country or territory	Indicator
Madagascar	77.8
DR Congo	77.1
Malawi	70.9
Guinea-Bissau	67.1
Zambia	64.4
Rwanda	60.4
Lesotho	59.7
Togo	54.2
Haiti	53.9
Benin	53.1
Sierra Leone	52.3
Tanzania	46.6
Niger	45.7
Burkina Faso	43.7
Chad	38.4
Senegal	38.0
Congo	37.0
Guinea	35.3
Uganda	34.6
Ethiopia	33.5
Sao Tome and Principe	32.3
Cameroon	24.0
Djibouti	22.5

Source: Prepared by the author based on the data published on https://data.worldbank.org/, accessed 14 April 2017.

The main characteristic of the BoP is the poverty of its population. In general, there is little information available on the composition of this population.

Education dimension

Many people and social groups have been incorporated into the education system in recent decades. Even many children from poor families now have the opportunity to enrol in school. The lack of schooling itself is a condition of exclusion. Today, not having that access increases marginality

and does not allow children to have the information they need to participate in society. But did this education expansion improve the relative position of poor families in the social structure? Or, did these changes allow intergenerational social mobility? There is no evidence that education systems are providing opportunities for intergenerational social mobility for children from poor families. But there is evidence that educational expansion has facilitated a certain intergenerational educational mobility. This is a result of public policies that seek to provide universal access to education, and the ability to finish it with a mastery of pre-academic skills that make it possible to continue learning throughout life.

Many children fail just when they begin their school life, especially those who do not have any kind of preschool education. In addition, many teachers who work with these children have not been prepared to address their specific problems.

There are several processes that contribute to this growing inequality. These processes operate within the school, outside of school, and in the interaction between the educational system and the broader social world. Within the system, one barrier is the need for initial preparation to start primary school, and the indirect costs of schooling (such as books, uniforms, etc.). Outside the education system, processes that intensify inequality arise from the living conditions of poor children, which weaken their health and make them more physically and psychologically vulnerable. All this discussion allows us to analyse the education dimension through three indicators: the rate of out-of-school children and adolescents of primary and lower secondary school age, adult literacy rate, and percentage of children who reach the end of primary education.

If we rank countries by the percentage of children and youth who are out of school, their positions change with respect to the poverty index, as shown in *Table 2.3*. The differences between one indicator and another show that wealth should not be the only indicator to measure poverty. This variable has data available for 124 countries. The maximum value is 61 per cent (Eritrea) and the minimum is 0.1 per cent (Japan). For this indicator, 32 countries are located in the highest quartile, at the BoP. About 263 million children and youth were out of school, according to data from the UNESCO Institute for Statistics (UIS) for the school year ending in 2014. This total includes 61 million children of primary school age, 60 million of lower secondary school age, and 142 million of upper secondary age (UNESCO, 2016). Several obstacles deprive children of their education. Most of these children and young people

Table 2.3 Rate of out-of-school children and adolescents of primary and lower secondary school age by country, highest quartile, latest available year between 2010 and 2015

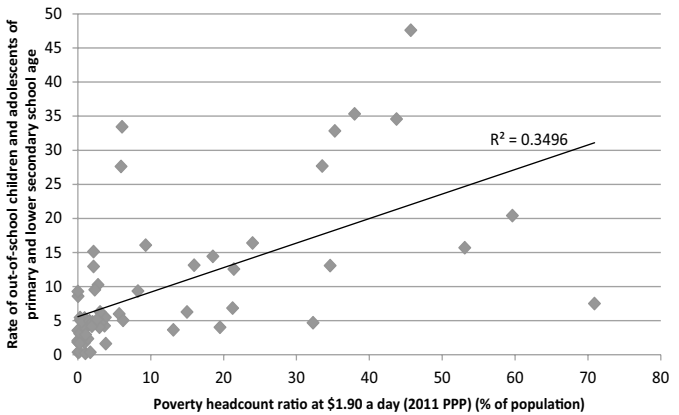
Country or territory	Indicator
Eritrea	61.0
Niger	47.6
Sudan	43.3
Mali	41.1
Central African Republic	39.1
Syrian Arab Republic	35.8
Senegal	35.3
Burkina Faso	34.6
Pakistan	33.4
Guinea	32.8
Ethiopia	27.7
Mauritania	27.6
Palau	23.7
Myanmar	22.8
Marshall Islands	21.7
Afghanistan	20.8
Lesotho	20.4
Mozambique	20.0
Comoros	19.5
Angola	18.2
Swaziland	18.0
Cameroon	16.4
Guatemala	16.1
Burundi	15.9
Benin	15.7
Cambodia	15.1
Puerto Rico	14.9
Lebanon	14.9
Bangladesh	14.4
Bermuda	14.2
Guyana	13.1
Honduras	13.1

Source: Prepared by the author based on the data published on http://data.uis.unesco.org/, accessed 14 April 2017.

live in countries affected by conflict or gender discrimination. The combination of exclusion and poverty causes children to enter the labour market early without the proper set of skills.

With respect to investment expansion and innovations, the high rates of out-of-school children and youth represent a present and future risk to BoP initiatives. This implies thinking about strategies to expand education systems. The following graph (*Figure 2.2*) shows a correlation between poverty and enrolment of the school-age population. In this group of countries, Malawi alone escapes the trend because 70 per cent of its population lives below the poverty line but less than 10 per cent of children and youth are out of school.

Figure 2.2 Poverty headcount ratio at $1.90 a day (2011 PPP) and rate of out-of-school children and adolescents of primary and lower secondary school age



Source: Prepared by the author.
Note: Data used is latest available between 2010 and 2015.

When individuals learn how to read, write, do basic math, and use computers, they have the power to lift themselves out of poverty, lower their health care costs, and find and keep sustainable employment. Literacy rates are available for 157 countries. The maximum value is 100 per cent (Uzbekistan) and the minimum is 19.1 per cent (Niger). When using this indicator, 39 countries are located in the lowest quartile (BoP).

An important condition for the success of the expansion of innovations and use of technologies is the user's ability to rapidly incorporate new knowledge. The focus on the quality of education has led to an emphasis on the measurement of learning outcomes (UIS, 2016). The results of the assessments are a plausible reference when

Table 2.4 Adult literacy rate, population 15+ years by country, lowest quartile, latest available year between 2010 and 2015

Country or territory	Indicator	Country or territory	Indicator
Niger	19.10	Guinea-Bissau	59.77
Guinea	30.47	Haiti	60.69
South Sudan	31.98	Bangladesh	61.49
Mali	33.07	Papua New Guinea	63.43
Central African Republic	36.75	Bhutan	63.91
Burkina Faso	37.75	Timor-Leste	64.07
Afghanistan	38.17	Madagascar	64.66
Benin	38.45	Nepal	64.66
Chad	40.02	Malawi	65.96
Côte d'Ivoire	43.27	Togo	66.54
Liberia	47.60	Yemen	69.96
Sierra Leone	48.43	Angola	71.16
Ethiopia	49.03	Rwanda	71.24
Mauritania	52.12	Morocco	71.71
Gambia	55.57	India	72.23
Senegal	55.62	Uganda	73.81
Pakistan	56.44	Eritrea	73.85
Sudan	58.60	Cameroon	74.99
Mozambique	58.84	Egypt	75.84
Nigeria	59.57		

Source: Prepared by the author based on the data published on http://data.uis.unesco.org/, accessed 14 April 2017.

analysing the competencies of the students and, consequently, also the functioning and efficiency of the education systems. Several countries with high levels of poverty do not participate in these types of evaluations due to the scarcity of resources. In this case, the percentage of children at the end of primary education achieving at least a minimum proficiency level is the relevant indicator. This variable has data available for 80 countries. The maximum value is 99.6 per cent (Netherlands) and the minimum is 8.5 per cent (Niger). For this indicator, 20 countries are located in the lowest quartile.

The OECD Programme for International Student Assessment (PISA) survey aims to evaluate the skills and knowledge of 15-year-old students. In our analysis we do not include this evaluation because almost all of these economies have either high or upper-middle incomes.

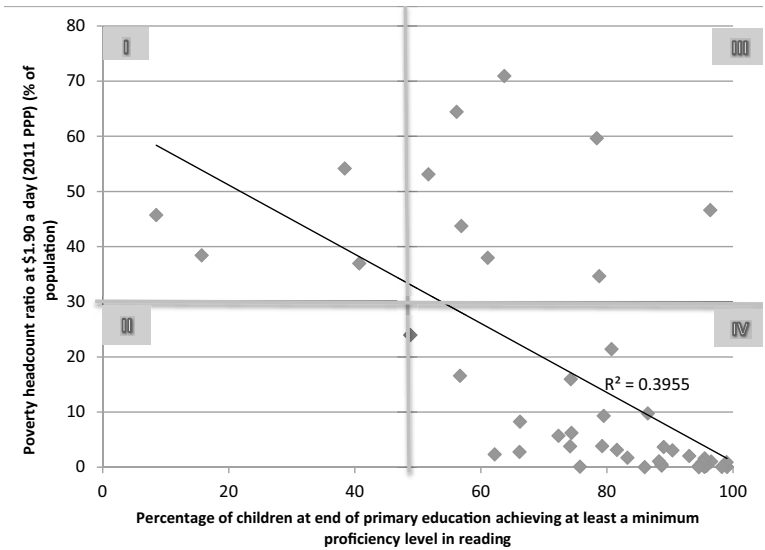
Table 2.5 Percentage of children at end of primary education achieving at least a minimum proficiency level by country, lowest quartile

Country or territory	Indicator
Niger	8.5
Chad	15.7
Morocco	21.0
Togo	38.4
Congo	40.7
Oman	47.4
Côte d'Ivoire	47.9
Cameroon	48.8
Benin	51.7
Botswana	55.7
Zambia	56.2
Burundi	56.5
South Africa	56.7
Burkina Faso	56.9
Kuwait	58.4
Qatar	59.7
Senegal	61.1
Dominican Republic	62.2
Malawi	63.7
United Arab Emirates	64.1

Source: Prepared by the author based on data on http://sdg4monitoring.uis.unesco.org/sdg4411, accessed 14 April 2017.
Note: Data used is latest available between 2009 and 2014.

In our analysis, 45 countries have a poverty index and learning outcomes data for reading at the end of primary education. Figure 2.3 shows the comparison between the two indicators. Generally, countries with high levels of poverty have high percentages of students who do not achieve the minimum necessary knowledge. Yet, in the group of countries in Quadrant III of the figure, the majority of students exceed the minimum levels of learning, despite high poverty levels. Of particular note is Tanzania, in which, despite high levels of poverty, the majority of students achieve at least the minimum proficiency level.

Figure 2.3 Percentage of children at end of primary education achieving at least a minimum proficiency level in reading against poverty headcount ratio



Source: Prepared by the author.
Note: Data used is latest available between 2009 and 2014.

Figure 2.3 presents four sections. The countries with high learning outcomes and low levels of poverty are concentrated in quadrant IV. Quadrant III has countries with high performance and high poverty levels. Quadrant II has only one country, in which half of its students do not reach the minimum levels of learning; Quadrant I, unsurprisingly, indicates the countries with high poverty and low reading achievement. To rapidly expand innovations and technologies, it is important that the population obtains at least the minimum level of competency. If we do not take into account learning outcomes, the projects and innovations would only cover part of the population of the BoP, or only some countries. At the same time, the expansion and survival of projects would be less likely to succeed.

Health dimension

Child mortality is a relevant indicator of the level of development in a country. It reflects the availability, utilization of, and access to health systems, particularly by mothers and children, as well as their nutritional status. This variable has data available for 194 countries. The maximum

value is 156.9 per 1,000 live births (Angola) and the minimum is 1.9 per 1,000 live births (Luxembourg). For this indicator, 48 countries are located in the highest quartile.

Table 2.6 Mortality rate, under-5 (per 1,000 live births) by country, highest quartile, 2015

Country or territory	Indicator	Country or territory	Indicator
Angola	156.9	Comoros	73.5
Chad	138.7	Zimbabwe	70.7
Somalia	136.8	Sudan	70.1
Central African Republic	130.1	Liberia	69.9
Sierra Leone	120.4	Haiti	69.0
Mali	114.7	Gambia	68.9
Nigeria	108.8	Lao PDR	66.7
Benin	99.5	Djibouti	65.3
DR Congo	98.3	Zambia	64.0
Niger	95.5	Malawi	64.0
Equatorial Guinea	94.1	Ghana	61.6
Guinea	93.7	Swaziland	60.7
South Sudan	92.6	Ethiopia	59.2
Côte d'Ivoire	92.6	Papua New Guinea	57.3
Guinea-Bissau	92.5	Kiribati	55.9
Afghanistan	91.1	Uganda	54.6
Lesotho	90.2	Timor-Leste	52.6
Burkina Faso	88.6	Turkmenistan	51.4
Cameroon	87.9	Gabon	50.8
Mauritania	84.7	Myanmar	50.0
Burundi	81.7	Madagascar	49.6
Pakistan	81.1	Kenya	49.4
Mozambique	78.5	Tanzania	48.7
Togo	78.4	India	47.7

Source: Prepared by the author based on the data published on https://data.worldbank.org/accessed 14 April 2017.

Information and communication technologies (ICT) dimension

The expansion of ICT has modified the tasks and the learning of families. But few studies have examined the abilities of the people who use these tools. In the ICT dimension, the percentage of individuals using a computer is an indicator that shows the level of exclusion from access to information and innovations. This variable has data available for 120 countries. The maximum value is 96.7 per cent (Iceland) and the minimum is 0.9 per cent (Burundi). For this indicator, 30 countries are located in the lowest quartile. In our analysis this group could represent the BoP.

Geographical context dimension

A country's wealth is also linked to the living conditions of its population. Countries in which a high percentage of the population lives in rural areas have difficulties providing basic services. While a few high-income countries have the majority of their populations in rural areas, almost all LICs have high percentages of their populations in rural areas without access to basic services.

In the last dimension, there is data available on rural population for 204 countries. The maximum value is 91.6 per cent (Trinidad and Tobago) and the minimum is 0 per cent (Singapore). For this indicator, 53 countries are located in the highest quartile.

The indicators analysed (see *Table 2.9*) show that we cannot identify the level of poverty or wealth of a country by looking at only one variable, or the context of its BoP status. If GDP per capita operates as an axis of the set of variables in our analysis, we could see which countries are in the most vulnerable quartile and, consequently, at the BoP. The countries with more appearances and low performances in our set of indicators will have a greater possibility of falling within the BoP category. Some countries were not included due to lack of data. On the resource side, tax revenues allow us to analyse the capacity of governments to cover the state's operating costs and mandatory services. In this group of countries, tax revenues are on average 19 per cent of GDP, the minimum is 0 per cent, and the maximum is 99 per cent.

This group of countries has a total of 928.8 million people. In 2015, the world poverty rate was 9.6 per cent. Yet, the number of poor remains high. 700 million people live on less than $1.90 a day – the updated international poverty line (World Bank Group, 2017). This population

is mainly concentrated in sub-Saharan Africa and South Asia. In the group of countries in *Table 2.9*, there are in total almost 280 million people living in poverty.

If we want to expand the base, we could include all countries where at least 1 per cent of their population is living in poverty. The following table shows countries with available information. In this group, 666.5 million people live in poverty. More than the exact number of countries that comprise the BoP, we should identify which ones at the BoP should be prioritized. The information in *Table 2.9* allows us to answer this question.

Some final thoughts

Processes that intensify inequality affect the living conditions of populations. Poor children are more likely to have health problems. For some poor families, short-term economic demands can make school attendance a luxury and not a regular part of life.

Context could be an additional factor, as well as the interaction between schools, homes, and children's expectations. These aspects all give rise to different educational climates for different children.

New policies are required to promote quality education, a contextualized educational model that can help problem-solving, and flexible planning that meets the new expectations and competencies of today's world. Fighting poverty is a need and an obligation. There will be little advancement if children cannot access quality education, and if we do not account for the clear relationship between development, education, and opportunities to escape poverty.

Table 2.7 Percentage of individuals using a computer by country, lowest quartile, latest available year between 2011 and 2015

Country or territory	Indicator
Burundi	0.9
Niger	1.5
Rwanda	2.0
Togo	2.3
Myanmar	3.1
Mali	3.3
Benin	3.6
Congo	4.0
Malawi	4.0
Lesotho	4.8
Mozambique	5.6
Bangladesh	5.7
Nepal	7.3
Senegal	8.0
Cambodia	9.3
India	9.5
Angola	10.4
Zimbabwe	10.7
Cameroon	11.8
Cuba	13.0
Sudan	14.0
Guatemala	15.2
Viet Nam	16.0
Bhutan	16.4
Kyrgyzstan	17.6
Indonesia	18.7
South Africa	20.8
El Salvador	22.3
Sri Lanka	22.4
Philippines	24.3

Source: Prepared by the author based on the data published on http://sdg4monitoring.uis.unesco.org/sdg4441, accessed 14 April 2017.

Table 2.8 Rural population (% of total population) by country, highest quartile, 2015

Country or territory	Indicator)	Country or territory	Indicator)
Trinidad and Tobago	91.6	Guyana	71.4
Burundi	87.9	Rwanda	71.2
Papua New Guinea	87.0	Burkina Faso	70.1
Liechtenstein	85.7	Channel Islands	68.5
Uganda	83.9	Barbados	68.5
Malawi	83.7	Tanzania	68.4
Sri Lanka	81.6	St Kitts and Nevis	68.0
St Lucia	81.5	Mozambique	67.8
Nepal	81.4	Zimbabwe	67.6
Niger	81.3	India	67.3
South Sudan	81.2	Timor-Leste	67.2
Samoa	80.9	Viet Nam	66.4
Ethiopia	80.5	Sudan	66.2
Cambodia	79.3	Myanmar	65.9
Swaziland	78.7	Bangladesh	65.7
Solomon Islands	77.7	Yemen	65.4
Micronesia, Fed. Sts.	77.6	Madagascar	64.9
Chad	77.5	Grenada	64.4
Tonga	76.3	Kyrgyz Republic	64.3
Antigua and Barbuda	76.2	Uzbekistan	63.6
Kenya	74.4	Guinea	62.8
Vanuatu	73.9	Lao PDR	61.4
Afghanistan	73.3	Bhutan	61.4
Tajikistan	73.2	Pakistan	61.2
Lesotho	72.7	Somalia	60.4
Comoros	71.7		

Source: Prepared by the author based on the data published on https://data.worldbank.org/, accessed 14 April 2017.

Table 2.9 Countries by number of available indicators of the dimensions of BoP

Country or territory	Region	Dimension 1	Dimension 2	Dimension 3			Dimension 4	Dimension 5	Dimension 6	Number of available indicators			
		GDP per capita	Poverty headcount ratio	Rate of Out-of-school children	Adult literacy rate	% of children achieving at least a min. proficiency level	Mortality rate, under-5 (per 1,000 live births)	% of individuals using a computer	Rural population (%)		Tax revenue (% of GDP) 2015	Population (Mill). 2015	Number of people living in poverty ($1.90 a day) (Mill)
Central African Repupblic	SSA	619	-	39.1	36.8	-	130.1	-	60.0	5	9.4	4.9	-
Burundi	SSA	727	-	15.9	85.5	56.5	81.7	0.9	87.9	7	-	11.2	-
DR Congo	SSA	784	77.1	-	77.2	-	98.3	-	57.5	5	8.8	77.3	59.6
Liberia	SSA	835	-	-	47.6	-	69.9	-	50.3	4	0.3	4.5	-
Niger	SSA	955	45.7	47.6	19.1	8.5	95.5	1.5	81.3	8	-	19.9	9.1
Malawi	SSA	1,184	70.9	7.5	66.0	63.7	64.0	4.0	83.7	8	15.2	17.2	12.2
Mozambique	SSA	1,192	-	20.0	58.8	78.0	78.5	5.6	67.8	7	23.1	28.0	-
Guinea	SSA	1,209	35.3	32.8	30.5	-	93.7	-	62.8	6	-	12.6	4.4
Guinea-Bissau	SSA	1,456	67.1	-	59.8	-	92.5	-	50.7	5	-	1.8	1.2
Togo	SSA	1,46	54.2	-	66.5	38.4	78.4	2.3	60.0	7	20.0	7.3	4.0
Madagascar	SSA	1,465	77.8	-	64.7	-	49.6	-	64.9	5	9.9	24.2	18.9
Comoros	SSA	1,483	-	19.5	78.1	-	73.5	-	71.7	5	-	0.8	-
Sierra Leone	SSA	1,569	52.3	-	48.4	-	120.4	-	60.1	5	8.6	6.5	3.4
Ethiopia	SSA	1,629	33.5	27.7	49.0	-	59.2	-	80.5	6	9.2	99.4	33.3
Gambia	SSA	1,68	-	-	55.6	-	68.9	-	40.4	4	-	2.0	-

Table 2.9, continued

Country or territory	Region	Dimension 1	Dimension 2	Dimension 3			Dimension 4	Dimension 5	Dimension 6	Number of available indicators			
		GDP per capita	Poverty headcount ratio	Rate of Out-of-school children	Adult literacy rate	% of children achieving at least a min. proficiency level	Mortality rate, under-5 (per 1,000 live births)	% of individuals using a computer	Rural population (%)		Tax revenue (% of GDP) 2015	Population (Mill). 2015	Number of people living in poverty ($1.90 a day) (Mill)
Burkina Faso	SSA	1,696	43.7	34.6	37.7	56.9	88.6	-	70.1	7	15.5	18.1	7.9
Haiti	LAC	1,757	53.9	-	60.7	-	69.0	-	41.4	5	-	10.7	5.8
Rwanda	SSA	1,762	60.4	-	71.2	-	41.7	2.0	71.2	6	14.0	11.6	7.0
Zimbabwe	SSA	1,787	21.4	12.6	86.9	80.7	70.7	10.7	67.6	8	-	15.6	3.3
Uganda	SSA	1,851	34.6	13.1	73.8	78.8	54.6	-	83.9	7	11.4	39.0	13.5
South Sudan	SSA	1,854	-	-	32.0	-	92.6	-	81.2	4	-	12.3	-
Afghanistan	SA	1,925	-	20.8	38.2	-	91.1	-	73.3	5	7.2	32.5	-
Kiribati	EAP	1,995	-	-	-	-	55.9	-	55.7	3	16.1	0.1	-
Mali	SSA	2,028	-	41.1	33.1	-	114.7	3.3	60.1	6	13.5	17.6	-
Benin	SSA	2,057	53.1	15.7	38.4	51.7	99.5	3.6	56.1	8	15.4	10.9	5.8
Chad	SSA	2,176	38.4	-	40.0	15.7	138.7	-	77.5	6	-	14.0	5.4
Solomon Islands	EAP	2,201	-	-	-	-	28.1	-	77.7	3	29.8	0.6	-
Timor-Leste	EAP	2,399	-	5.4	64.1	-	52.6	-	67.2	5	99.9	1.2	-
Senegal	SSA	2,421	38.0	35.3	55.6	61.1	47.2	8.0	56.3	8	19.0	15.1	5.7
Nepal	SA	2,462	15.0	6.3	64.7	-	35.8	7.3	81.4	7	16.7	28.5	4.3
Tanzania	SSA	2,673	46.6	-	80.4	96.4	48.7	-	68.4	6	12.4	53.5	24.9
Yemen	MENA	2,821	-	-	70.0	-	41.9	-	65.4	4	-	26.8	-

Table 2.9, continued

Country or territory	Region	Dimension 1	Dimension 2	Dimension 3			Dimension 4	Dimension 5	Dimension 6	Number of available indicators			
		GDP per capita	Poverty headcount ratio	Rate of Out-of-school children	Adult literacy rate	% of children achieving at least a min. proficiency level	Mortality rate, under-5 (per 1,000 live births)	% of individuals using a computer	Rural population (%)		Tax revenue (% of GDP) 2015	Population (Mill). 2015	Number of people living in poverty ($1.90 a day) (Mill)
Tajikistan	ECA	2,834	19.5	4.0	99.8	-	44.8	-	73.2	6	-	8.5	1.7
Lesotho	SSA	2,95	59.7	20.4	79.4	78.4	90.2	4.8	72.7	8	49.1	2.1	1.3
Vanuatu	EAP	2,988	15.4	-	85.1	-	27.5	-	73.9	5	16.0	0.3	0.0
Kenya	SSA	3,089	-	10.7	78.0	92.1	49.4	-	74.4	6	15.5	46.1	-
Cameroon	SSA	3,115	24.0	16.4	75.0	48.8	87.9	11.8	45.6	8	-	23.3	5.6
Sao Tome and Principe	SSA	3,219	32.3	4.7	91.7	-	47.3	-	34.9	6	14.6	0.2	0.1
Bangladesh	SA	3,34	18.5	14.4	61.5	-	37.6	5.7	65.7	7	9.0	161.0	29.8
Kyrgyzstan	ECA	3,434	1.3	-	99.5	-	21.3	17.6	64.3	6	17.7	6.0	0.1
Cambodia	EAP	3,49	2.2	15.1	78.3	-	28.7	9.3	79.3	7	14.6	15.6	0.3
Djibouti	MENA	3,491	22.5	-	-	-	65.3	-	22.7	4	-	0.9	0.2
Micronesia, FS	EAP	3,497	17.4	-	-	-	34.7	-	77.6	4	12.6	0.1	0.0
Côte d'Ivoire	SSA	3,514	-	-	43.3	47.9	92.6	-	45.8	5	14.4	22.7	-
Zambia	SSA	3,836	64.4	-	85.1	56.2	64.0	-	59.1	6	16.1	16.2	10.4
Marshall Islands	EAP	3,911	-	21.7	98.3	-	36.0	-	27.3	5	-	0.1	-

Source: Prepared by the author.

Note: East Asia Pacific (EAP), Europe and Central Asia (ECA) Latin America and Caribbean (LAC), Middle East and N. Africa (MENA), South Asia (SA) and sub-Saharan African (SSA).

Table 2.10 Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population). Last year available between 2011 and 2015

Name of country	Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population)	Number of people living in poverty (Mill)	Name of country	Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population)	Number of people living in poverty (Mill)
Madagascar	77.8	18.9	Georgia	9.8	0.4
DR Congo	77.1	59.6	Guatemala	9.3	1.5
Malawi	70.9	12.2	Indonesia	8.3	21.3
Guinea-Bissau	67.1	1.2	Bolivia	6.8	0.7
Zambia	64.4	10.4	Nicaragua	6.2	0.4
Rwanda	60.4	7.0	Pakistan	6.1	11.5
Lesotho	59.7	1.3	Mauritania	5.9	0.2
Togo	54.2	4.0	Colombia	5.7	2.7
Haiti	53.9	5.8	Ecuador	3.8	0.6
Benin	53.1	5.8	Panama	3.8	0.2
Sierra Leone	52.3	3.4	Brazil	3.7	7.6
Tanzania	46.6	24.9	Peru	3.1	1.0
Niger	45.7	9.1	Viet Nam	3.1	2.8
Burkina Faso	43.7	7.9	Mexico	3.0	3.9
Chad	38.4	5.4	El Salvador	3.0	0.2
Senegal	38.0	5.8	Paraguay	2.8	0.2
Congo	37.0	1.7	Tuvalu	2.7	0.0
Guinea	35.3	4.5	Dominican Republic	2.3	0.2

Uganda	34.6	13.5	Armenia	2.3	0.1
Ethiopia	33.5	33.3	Bhutan	2.2	0.1
Sao Tome and Principe	32.3	0.1	Cambodia	2.2	0.3
Cameroon	24.0	5.6	Bulgaria	2.0	0.2
Djibouti	22.5	0.2	Tunisia	2.0	0.2
Zimbabwe	21.4	3.3	Sri Lanka	1.9	0.4
India	21.2	278.3	China	1.9	25.4
Tajikistan	19.5	1.7	Argentina	1.7	0.7
Bangladesh	18.5	29.8	Costa Rica	1.6	0.1
Micronesia, Fed. Sts.	17.4	0.1	Latvia	1.4	0.1
Lao PDR	16.7	1.1	Kyrgyz Republic	1.3	0.1
South Africa	16.6	9.1	Seychelles	1.1	0.0
Honduras	16.0	1.3	Albania	1.1	0.1
Vanuatu	15.4	0.1	Lithuania	1.0	0.1
Nepal	15.0	4.3	Estonia	1.0	0.0
Philippines	13.1	13.2	**TOTAL**	-	**666.5**

Source: Prepared by the author based on the data published on https://data.worldbank.org/, accessed 14 April 2017.

References

Dubois, A. 2009. 'Wealth redistribution'. In: *Education for sustainability manual*, no. 13. Bilbao, Spain: HEGOA Institute, University of the Basque Country. Retrieved from: www.unescoetxea.org/ext/manual_EDS/pdf/13_riqueza_castellano.pdf

Prahalad, C.K. 2005. *The fortune at the bottom of the pyramid*. New York: Pearson Education. Retrieved from: http://etelangana.org/uploads/ebooksCkPrahalad_2014-03-26_130317/2014-03-26_130520_Fortune.pdf

UNESCO. 2016. 'Leaving no one behind: how far on the way to universal primary and secondary education?' Institute for Statistics and Global Education Monitoring Report Policy paper 27/Fact sheet 37. Retrieved from: http://uis.unesco.org/sites/ default/files/documents/fs37-leaving-no-one-behind-how-far-on-the-way-to-universal-primary-and-secondary-education-2016-en. pdf

UIS (UNESCO Institute for Statistics). 2016. 'The cost of not assessing learning outcomes', Information Paper no. 26. Retrieved from: http://uis.unesco.org/sites/default/files/documents/the-cost-of-not-assessing-learning-outcomes-2016-en_0.pdf

——. UIS Database. Accessed 14 April 2017: http://data.uis.unesco.org/

——. Monitoring SDG4 Database, accessed 14 April 2017: http://sdg4monitoring.uis.unesco.org/sdg4111

——. Monitoring SDG4 Database, accessed 14 April 2017: http://sdg4monitoring.uis.unesco.org/sdg4441

World Bank. Database. Accessed 14 April 2017: https://data.worldbank.org/

Commentary

Justin Sandefur

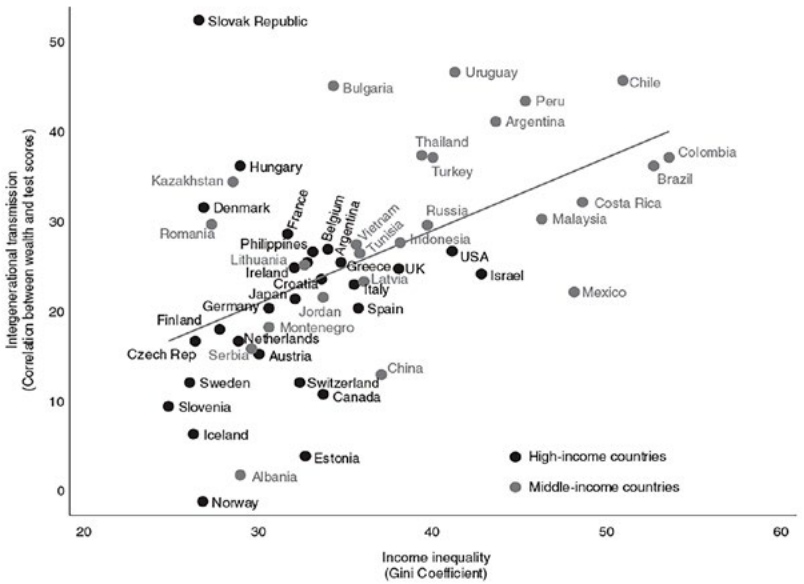
In her analysis of the bottom of the pyramid in global education, Dr Montoya focuses quite sensibly on low-income countries, documenting the characteristics and development challenges of this group, and their lagging education outcomes. Implicitly, this is a decision to focus on inequality between countries rather than within countries. Three points should be made: two suggesting this is the right choice, and one in favour of greater focus on domestic inequality and class differences within a given developing country.

1. *Global income inequality is mostly driven by differences between rich and poor countries, not the differences between rich and poor within a given country*. This simple observation is the conclusion of Milanovic's (2015) analysis of income and expenditure surveys from around the world. Decomposing income inequality into the within- and between-country components, he finds that the latter accounts for just over half. Furthermore, this appears to be a modern phenomenon. The age of class differences has gradually faded in relative importance, and country of birth now matters much more in explaining your income.
2. *Learning levels in rich and poor countries diverge widely*. Anyone who has glanced at the results from major international learning assessments knows that pupils in poor countries perform well below their rich-country counterparts. These differences are rather extreme. In the case of PISA for instance, only about 1 per cent of South African children achieve Singapore's average PISA score. So global inequality in education is largely about the failure of entire education systems, like South Africa's, not just inequality within them.
3. *On the domestic front, as within-country income inequality rises, the learning gap between rich and poor widens even faster*. Education is often billed as a ladder out of poverty. But looking again at PISA scores, we see that this ladder begins to crumble as economic inequality grows. In countries with high income inequality, the relationship between household wealth and student performance on learning assessments becomes much stronger.

This general relationship, in which high income inequality is associated with less economic mobility across generations, is commonly

referred to as 'The Great Gatsby curve'. The curve is particularly striking when using test scores as a measure of outcomes. For instance, if your household is one standard deviation richer than your neighbour's in Brazil, you can expect to score about 0.3 standard deviations higher on the PISA. If you have the same wealth gap in Canada, you get basically no advantage in test scores (Sandefur, 2015).

Figure 2.4 Income inequality and intergenerational transmission in high- and middle-income countries, 2012



Source: Author's calculations based on PISA 2012 microdata.

The Gatsby curve implies the poor are doubly damned. Not only do rich people use their wealth to buy better education for their kids, but as inequality increases, families with greater wealth pass on more and more of that advantage to their children.

In sum, should global education policy debates prioritize inequality within or between countries? That's too big a question to settle here. But in the short term, feasibility matters. The education policy literature suggests that we don't know how to make South Africa into Singapore, but we do know how to educate poor Brazilian kids like rich Brazilian kids, so perhaps that's a good place to start.

References

Milanovic, B. 2015. 'Global inequality of opportunity: How much of our income is determined by where we live?' In: Review of Economics and Statistics, 97(2), 452–460.

Sandefur, J. 2015. 'Great Gatsby revisited: how inequality explains learning outcomes around the world'. Center for Global Development. Retrieved from: www.cgdev.org/blog/great-gatsby-curve-younger-and-poorer-how-inequality-explains-learning-outcomes-around-world

Commentary

Amrit Thapa

The paper by Dr Silvia Montoya explores the understanding of the term 'bottom of the pyramid' (BoP) in the case of low-income countries. The paper proposes analysing six dimensions in order to better understand the countries at the BoP, rather than just using GDP per capita. This is premised on the argument that one needs to take into account other socio-economic and geographic dimensions for a better picture of the problems these countries are facing. These six dimensions are: wealth, poverty, education, health, ICT, and geographical context.

The paper, which presents interesting statistics and graphs, is very illuminating in its proposition to obtain a more robust definition of BoP by including a variety of important dimensions, rather only one or two indicators. It also provides valuable insights into how we can better define poverty and related concepts. However, there are several ideas worthy of further reflection.

First, the six dimensions used in the paper are well thought out in tapping into the major dimensions of a country's performance. However, it would be helpful if the paper had included indicators on 'corruption' and 'governance' too. A substantial amount of foreign aid is being poured into these countries that are at the BoP. Unfortunately, due to poor governance or corruption, much of these funds does not reach the places or people who actually need them. As a result, the development process is badly hindered, and the countries continue to be trapped in the poverty cycles. Therefore, if not as part of the dimensions measured, at least some discussion on these important topics of governance and corruption would have been meaningful.

The other major challenge is that the countries that are at the BoP are far behind in terms of data collection, and the availability of some essential indicators for education development. These countries lack well-established systems that integrate different data sources. As a result, the monitoring of SDGs in these countries is hindered. The biggest gaps in data availability concern direct assessment of ICT skills or digital literacy, global citizenship and education for sustainable development, scholarships, and development outcomes of young children. The UIS report indicates that nearly all countries collect data on students according to sex and whether living in urban or rural areas. However, data on children with disabilities or special learning needs

are collected in only 62 per cent of countries, and other population groups (e.g. refugees, out-of-school children, nomadic groups, orphans and vulnerable children) are missing entirely in most national statistics (UIS, 2016). In this context, without good and sufficient data, it can be challenging and at times misleading to attempt to understand a country's performance and rank in terms of development indicators. Perhaps further effort should be devoted to this approach and how best to address these limitations.

Finally, in the education sector, the assessment and comparison of data relating to learning and skills is a complicated task. For example, despite a growing number of learning assessments, there is currently no mechanism to reconcile the differences between various types of assessment to generate cross-nationally comparable data. Moreover, when it comes to countries at the BoP, as previously mentioned, a host of other issues such as corruption, political instability, poor infrastructure and other technical problems add to the challenge. Unified and persistent efforts from national and international communities are needed to overcome these challenges, and to make the educational development efforts in low-income countries more effective.

References

UIS (UNESCO Institute for Statistics). 2016. *Sustainable development data digest: laying the foundation to measure Sustainable Development Goal 4*. UNESCO Institute of Statistics, Montreal, Canada. Retrieved from: http://uis.unesco.org/sites/default/files/documents/laying-the-foundation-to-measure-sdg4-sustainable-development-data-digest-2016-en.pdf

Chapter 3
Making the pyramid less pyramidal?

Luis Crouch

Introduction

Business literature on the 'bottom of the pyramid' claims that countries and businesses can achieve great progress and growth by both augmenting and catering to the purchasing power of the poor (Pralahad, 2005). Analogizing from this notion, Wagner has called for a similar focus on improving learning outcomes at the bottom of the pyramid (Wagner, 2011; Wagner and Castillo, 2014; Introduction of this volume). His central claim is that education policies and approaches need to work with the poorest (or those with lowest learning achievement as proxied by various assessments), which may require a different approach from that used with the middle or the top of the pyramid, namely the middle classes in middle- and higher-income countries.

In this chapter, we bring empirical evidence to bear on two points related to Wagner's thesis. First, we briefly investigate whether cognitive performance is indeed pyramidal, at least in the countries with the lowest cognitive performance. Second, we explore whether the fastest overall progress can be made by working at the bottom, that is, in going from the bottom to the middle of the cognitive pyramid, rather than going from middle to top.

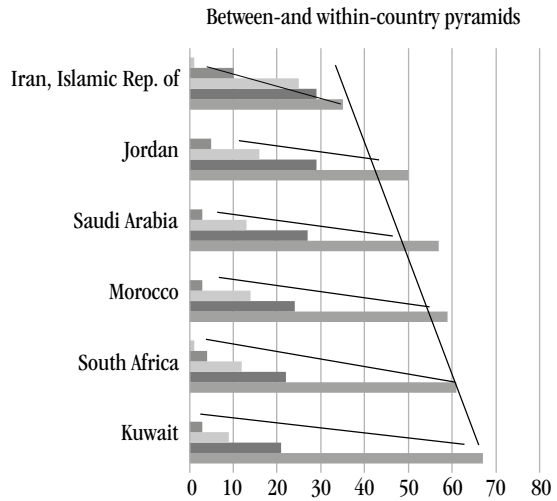
Is the cognitive distribution pyramidal?

Judging from the graphical shape of the distribution of cognitive (learning) achievement in the countries with relatively low average cognitive performance, the answer is a clear 'yes'. However, because the implicit sampling of the main international assessments is not proportional to population (since learning outcomes data on the entire population of very large developing countries are not known with any degree of certainty, even if data for some cities or sub-national entities are known), it is difficult to assess how pyramidal the global distribution is.

However, when comparing *individual* countries that are towards the bottom of the performance distribution, it is clear that the distribution of cognitive skills is pyramidal in nature. Moreover, this distribution

is pyramidal both between *and* within countries. *Figure 3.1* makes this clear, using the Trends in International Mathematics and Science Study (TIMSS) 2015 Fourth Grade Mathematics. The countries are presented, from bottom to top, in order of *increasing* average scores on the assessment. The horizontal bars for each country represent the proportion of the assessed students for each of the five TIMSS levels of proficiency, from the lowest levels at the bottom of each country's pyramid, to the higher levels towards the top of each pyramid. (For most countries the fifth – highest – level is so small that the bars do not show up. Iran and South Africa are the exceptions.) A similar structure can be found using the Progress in International Reading Literacy Study (PIRLS) 2011 (Mullis *et al.*, 2012) or PISA 2015 (OECD, 2016). The lines superimposed on the bars are a heuristic for the edges of the pyramids, first for the overall 'between-county' pyramid (the lines tracing the *bottom* of each country's pyramid), and second for each country's individual pyramid. Naturally, the figure shows only one 'edge' of the pyramid. The conclusion is that, using the noted TIMSS results, there is a clear pyramidal structure both within and between countries.[6]

Figure 3.1 Pyramidal nature of learning distribution in lower-middle-income countries



Source: Prepared by author based on TIMSS 2015 Grade 4 Mathematics data, at: http://timss2015.org/download-center/.

6. Both TIMSS and PIRLS show the most pyramid-like structure, relative to PISA. And, TIMSS data (in particular) suggest that as countries progress to the highest level, the pyramid, if anything, becomes inverted. (See figures below.) Thus, one has to take care with the pyramid figure of speech, useful as it may be.

Is faster progress possible by working from the bottom?

Typically, education policy-makers are interested in improving learning outcomes. But the highest-level policy-makers outside the field of education – those in cabinet, or the legislature, or the ministry of finance – tend, naturally, to focus on a single indicator, such as the average test score (be it on an international or a national assessment), or the percentage of students above a certain threshold. This is understandable: policy-makers have limited scope of control and attention, and the publics do as well. However, educators' and economists' goals tend to emphasize distributional issues in addition to average ones. That is, they are interested in improvements along the whole distribution, along with increases in the total or the average. The SDGs, even if many economists decry them as 'not quite what they'd like', will galvanize attention, and their focus is squarely on equality, not just on averages. Words such as 'for all,' 'inclusive', 'equal', 'equality', 'equity', and 'equitable' are liberally sprinkled throughout the discussion of the education goal, and the indicators all ask for disaggregation along various dimensions of equality. Making an analogy to economic growth, one can work to raise the overall or average per capita rate of growth of an economy, and also work to make progress on both absolute and relative poverty. Furthermore, in focusing on poverty, looking at a single indicator, such as the percentage of the population living under a given 'dollar per day' benchmark, has been criticized as too simplistic (e.g. see Pritchett, 2013). So, even when looking at poverty (income poverty, or poverty in cognitive and non-cognitive or other skills), it pays to look at whole distributions, not just single benchmarks.

So, we seem to have a bit of a dilemma. High-level policy-makers and the public typically like single indicators that point to a central tendency. Furthermore, if they think about it at all, these policy-makers often think that the way to improve the average is to focus on those at the middle of the distribution, or even at the upper end: that the 'right hand' of the distribution can 'pull over' the whole distribution. Thus, a common policy outcome in many low- or middle-income countries is to over-invest in free tertiary education and under-invest in early childhood development for the poor, or, to create magnet schools or schools for the talented. Of course, this lines up with the political economic incentive of catering to vocal urban voters, those who are 'destined' to go on to tertiary education, and so on. Expenditure on these social groups is higher, and, though less fashionable than they once were, 'magnet' or 'model' schools are not uncommon, and spending and attention devoted

to the poor would be lower were it not for the pressure created by goals such as the SDGs. Professionals on the other side, such as educators and economists, tend to want to look at the whole distribution, and to look at the distribution in at least two ways: one that focuses on absolute poverty, and one that focuses on relative poverty.

But maybe there is a relatively simple way out of this dilemma or this conceptual gap between educators and economists on the one hand, and higher-level, non-sectoral policy-makers on the other – at least in terms of conceptual focus. Perhaps countries that succeed at bringing up their averages or means, from the lowest levels to the middling levels, do it not so much by 'pulling' the distribution to the right from the left, but by 'pushing' it to the right from the left. In that sense, what if reducing *absolute* cognitive poverty, starting with the lowest possible levels, could *also* reduce relative cognitive poverty, and was the easiest way to improve the average?

This section will argue that there may be reasons to have hope in this respect, and that a relentless focus on the achievement of those who achieve the least, whoever they may be, is a way to generate economy of attention and economy of effort. We will not argue that the task is thereby made any easier, however, in terms of technical policy design, implementation, or political economy.

How do countries actually make progress at the average?

This section explores the idea that countries make more progress in increasing average performance by reducing the number of students in the lowest proficiency levels than by increasing the number of students in the highest proficiency levels. To explore this, we take advantage of an empirical regularity that is implicit in the results of international assessments. We have taken three assessments as cases-in-point: PISA Reading 2015 (OECD, 2016), PIRLS 2011 (Mullis *et al.*, 2016), and TIMSS Mathematics Grade 4 (see *Figure 3.1*). Further research could explore the pattern in other assessments.

The relevant data are the relative frequency distributions of performance at the various cut-off points of the proficiency scales for these assessments, including the bottom one of 'less than X'. As a reminder, note that these proficiency scales have at least some cardinal value. That is, they are designed so that moving from level 2 to level 3

is considered as difficult as moving from level 3 to 4, or 4 to 5. They are 'real metrics' in this sense, not just ordinal scales.[7]

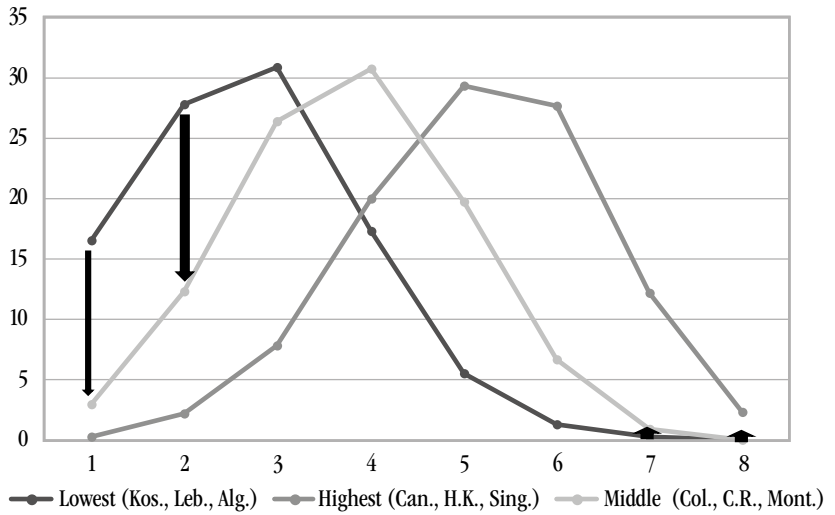
The exploration of this issue, while not common in the literature, is not unique to this paper. Van Damme (2017) notes, for instance, that the correlation between the overall PISA 2015 science score and the distance between performance at the median and performance at the 10th percentile is 0.56. However for reading, he finds only 0.14, and (with respect to reading) he seems to come to a different conclusion than ours, though he uses a different method. His method looks at the correlation between gaps in achievement at the 10th and 50th percentiles and average country performance (at a given point in time), but does not look at the proportions of the student population that are at different levels of proficiency, and how those proportions differ according to average level of performance. Mullis and colleagues look at the long-term dynamics (1995–2015), not a given point in time as we have done, using TIMSS Mathematics and Science, Grade 4 for various years (Mullis *et al.*, 2016).[8] For mathematics, they note that it is more common for countries to improve by making gains at the 10th percentile of their results distribution as opposed to higher up: 'More gains at the 10th percentile is the dominant pattern' (Mullis *et al.*, 2016: 59). This is particularly true for countries that improved their average scores the most (more than 10 per cent) and were below the median in 1995. For these four countries, there was a 71 per cent greater improvement at the 10th percentile than at the 90th percentile. The pattern is similar in science. They note further: 'The five countries with the biggest TIMSS gains over time – Slovenia, Singapore, Portugal, Hong Kong SAR, and the Russian Federation – all registered larger gains at the 10th percentile than the 90th percentile' (Mullis *et al.*, 2016: 59).

To make the key tendencies visible, the graphics below, starting with *Figure 3.2*, show the cumulative frequency distributions, over levels of proficiency, for three groups of countries: a set of high performers, a set of middling performers, and a set of low performers, *defined on the average or median performance*. Thus, the lowest group

7. This is implicit in the method for calculating the proficiency scales in these assessments, which, to over-simplify, starts by sorting students from least to most able according to the percentage of items they answer correctly. Note that not all items are of equal difficulty, and the more difficult items may 'weigh' more, which might lead to a re-sorting of the students (i.e. someone who answered fewer items correctly, but who got correctly items that were far more difficult than those answered by others, may now 'leap' over other students in the re-sorting).
8. The disadvantage of looking at long-term dynamics is that there is only a reduced sample of countries that participated in TIMSS in both 1995 and 2015.

is the lowest in terms of their performance at the 50th percentile of their own distribution. The figures demonstrate that, in moving from 'worst' to 'middle' performance, the decline in the lowest two levels of performance (in PISA) and in the single lowest level of performance in TIMSS and PIRLS is much greater than the improvement in the highest two levels of performance in PISA or the highest single level of performance in TIMSS and PIRLS. This analysis assumes that a cross-section of countries' performance is a reasonable proxy for dynamic development of particular countries' performance.

Figure 3.2 Distribution of reading performance in PISA 2015 (% of students by proficiency level)



Source: Calculated by the author based on OECD, 2016.

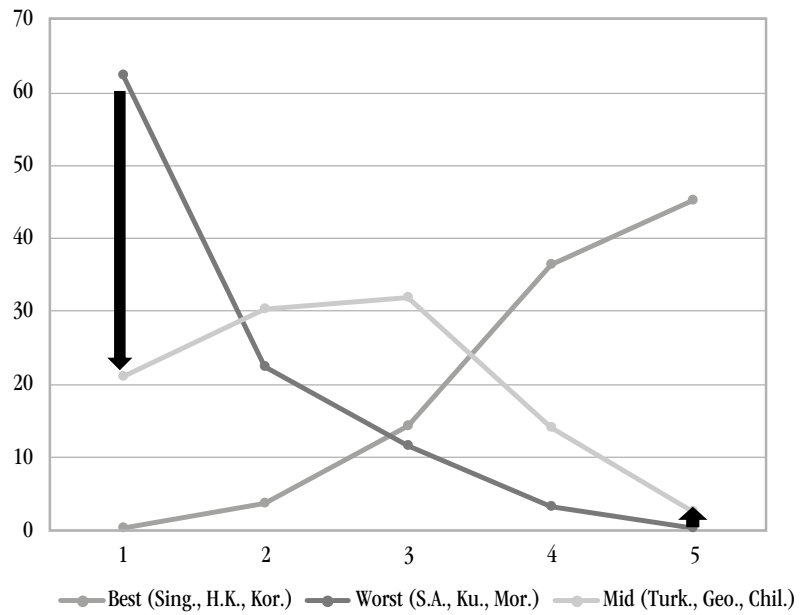
Figure 3.2 for PISA 2015 (reading) shows a substantial decline in the lowest two rungs of the proficiency distribution, in moving between the lowest performers and the highest performers, and only a tiny increase in the top two rungs. In going from the middle-performing levels to the high levels, there is still a decline in the lowest rungs, but there is a relatively large increase in the top rungs, in particular the second rung from the top.

Similar analyses were created for PISA 2015 science and mathematics, but the results do not vary significantly. (PISA 2015 science deviates slightly from the patterns of all other assessments noted in this paper: in the lowest-performing countries, the lowest proficiency category does

not contain the most students. In that sense, it is the only assessment that is not quite 'pyramidal' among the relatively low-performing countries.) The results for PISA 2015 science and mathematics, though not presented here graphically, are tabulated in *Table 3.1*, which contains results from all assessments discussed in this paper.

The data for TIMSS are clearer and more dramatic, perhaps because TIMSS uses fewer proficiency categories. With fewer proficiency categories, the proportions of students falling into those categories are larger and less subject to irregularities.

Figure 3.3 Distribution of Grade 4 mathematics performance in TIMSS 2015 (% of students by proficiency level)

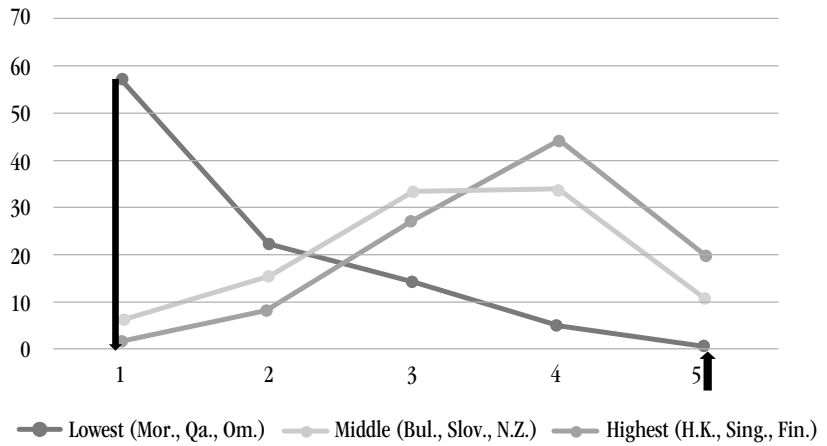


Source: Calculated by the author from TIMSS Database, IEA's Trends in International Mathematics and Science Study – TIMSS 2015.

Finally, for PIRLS 2011 (*Figure 3.4*), the results are more or less the same, though the curves behave a little differently.

Graphs like these can be misleading, because they show only the frequencies. Yet the mean value, or expected value, over the whole

Figure 3.4 Distribution of performance in PIRLS 2011 (% of students by proficiency level)



Source: Mullis *et al*. 2012: 68–69.

distribution is the sum of the product of frequencies times the average level of performance *within* each proficiency interval in the frequency distribution. And those values are higher towards the right-hand side. So, the graphs probably create an exaggerated visual impression. Relatedly, given how skewed these distributions are, it is hard to discern the mean and median of the distribution graphically, as they are clearly very far from the mode. Thus, it seemed wise to quantify numerical simulations in addition to providing the graphical illustrations.

First, the mean or expected PISA, TIMSS, or PIRLS score for any given country is the sum, across all proficiency levels, of the average level of performance (or the average proficiency) *within* each proficiency level, times the percentage of students falling into that proficiency level. This is based on the standard definition of a weighted average or an expected value. That is, each country's performance is simply the weighted average of the performances at each proficiency level, where the weights are the proportions of the students at those levels of performance.

One has to acknowledge that the average level of proficiency within each proficiency level varies by country (because there is varying skewness between countries even within proficiency categories), but that this variation is likely to be small, and, especially, that the *difference* in the typical proficiency levels within proficiency categories *within country* (that

is, the differences in within-category skewness in any given country) and *over time* (using countries to proxy time) are likely to be very small.

Second, we say that for a typical country moving up the performance scale, the *change* in average performance is the change in the proportion of students at each level of performance times the fixed level of performance in each proficiency level. It is assumed that the average performance *within* proficiency categories does not change.

Also, it is important to quantify how much the change in the proportion of students at each level of performance contributes to the change in average performance. This way, we can quantify how much the *reduction* in the proportion of students at the bottom of the pyramid contributes to improving the average, versus how much the *increase* in the proportion of students at the top of the pyramid contributes to the increase.[9]

For PISA, the results were that the contribution of reducing the lowest two levels to the change in the mean between the bottom and the middle countries, was 93 points. The contribution to the change in mean of increasing the relative frequency at the top two levels was 45 points. The reduction in relative frequency in the bottom was twice as important as at the top. When going from middle to upper, the results flip, and are less dramatic: a contribution of 45 points via the increase in the relative frequency in the top two levels, and a contribution of only 18 points via reducing the relative frequency in the bottom two levels.

For TIMSS, the present analysis focuses only on the very bottom and top categories, since there are fewer categories in TIMSS. Taking the top

9. To prevent reasoning from outliers, we took the average proficiency proportions in the bottom five countries, the five countries in the middle, and the top five countries, as 'typical' of the bottom, middle, and top of the pyramid. This naturally works a little against our own hypothesis (because it reduces the extremes), but it seems safe to prevent the influence of outliers and irregularities in the data.

One final technical detail is that because of how the data are reported, the open-ended (above the top cut-off point, and below the bottom cut-off point) categories show no average performance, so one has to impute. For PISA, we took the mid-point of the distance between the cut-off points, and then assumed that the mean within-category proficiency for the two open-ended categories was as far below or above the cut-off point as the cut-off point was from the next cut-off point. This fits the data well, in that the expected value using the frequency distributions of the proficiency levels, and the mid-points and extended points for the open categories, was quite close to the mean performance for the bottom five, middle five, and top five country groupings: a mean absolute difference of only 1, in a scale with a mean of around 500. For TIMSS and PIRLS the procedure did not fit as well, because the bottom group is so large among the poorer performers. For the top level we extrapolated the average distance exactly as for PISA. For the bottom level we extrapolated more: enough to drive the mean absolute deviation of the estimated country group averages to a minimum (3.5 for TIMSS and 4.9 for PIRLS out of a scale centred at 500, so a very small approximation error), using a simple linear optimization algorithm.

two and the bottom two would mean taking 80 per cent of the categories, which would defeat the purpose. The results show that in going from the poorest overall performance to the median overall performance, reducing the size of the relative frequency at the bottom level contributes 162 points, and increasing the size of the relative frequency at the top level contributes only 46 points. In going from median overall performance to top overall performance, reducing the size of the bottom proficiency frequency contributes only 17 points, but increasing the size of the top proficiency frequency contributes 221 points.

Finally, for PIRLS, the reduction in the bottom category contributes 133 points to the increase in the average, whereas the increase in the top category contributes only 64 points, in going from the bottom to the middle. The continued reduction in the bottom category contributes 11 points to further increasing the average, whereas the increase in the top category contributes 62 to the increase in the average.

Table 3.1 Summary of the results of all three assessments

Assessment	Moving from lowest country average performance levels to middle country average performance levels		Moving from middle country average performance levels to high country average performance levels	
	Reduction in numbers in lowest proficiency levels	Increase in numbers in highest proficiency levels	Reduction in numbers in lowest proficiency levels	Increase in numbers in highest proficiency levels
PISA 2015 Reading	93	45	18	45
PISA 2015 Science	79	26	12	94
PISA 2015 Mathematics	189	58	71	123
TIMSS 2015, Mathematics Grade 4	162	46	17	221
PIRLS 2011	133	64	11	62

Sources: Calculated by the author based on data from OECD, 2016. For TIMSS, TIMSS Database, IEA's Trends in International Mathematics and Science Study – TIMSS, 2015, http://timss2015.org/download-center/. For PIRLS, Mullis *et al.* 2012.

The results hold across these three assessments (five, if one counts the PISA variants): in going from low to middle levels of average country performance, the impact of reducing the proportion of students at the lowest proficiency level is much greater than the impact of increasing

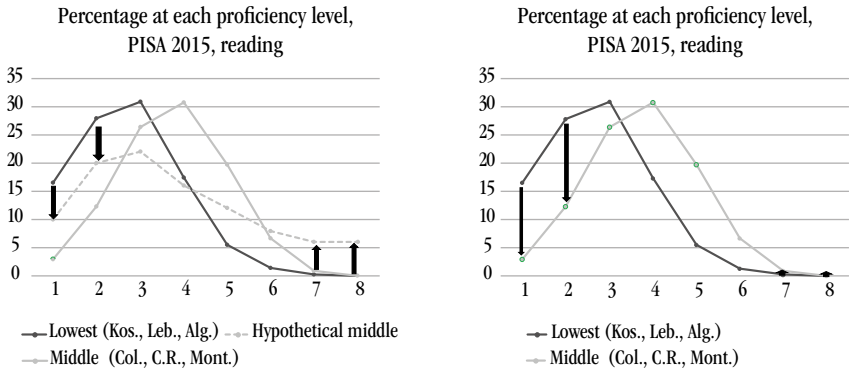
the proportion of students at the highest proficiency level. In going from the middle average country performance to highest average country performance, the impacts are reversed.

The results would be trivial if what is shown above *has* to be true, by definition or mathematical artifact. But there is no reason why the results simply have to be true by definition. The results could also be trivial if, as some tend to think, the phenomenon is simply what happens as nations evolve, and has nothing to do with purposeful (education) policy. Part of the reason, for instance, could be that the top performers top out: one cannot move further up. We can refute the first form of possible triviality in two ways. First, via a counter-example, and, second, by noting that even in the best-performing countries, only an extremely tiny minority of students gain a score of 100 per cent.

A numerical counterexample is provided for PISA showing the same (real situation) graphic as shown above (with the highest level removed, for clarity), on the right, and a hypothetical one on the left (*Figure 3.5*). The hypothetical case seems quite reasonable and plausible. It produces an increase in the mean performance level for the middle countries equal to the real case. And it yields that increase through a reduction in the bottom two proficiency levels, which contributes only 38 points to the increase in the mean, while the increase in the top two proficiency levels contributes 82 points: the opposite of what happens in reality, but with the same net effect. Note, however, that in the hypothetical graphic the two arrows on the right are of the same size as the two arrows on the left (more or less), but the underlying within-group proficiency levels are much higher for the two arrows on the right. That is what allows the effect from the two right arrows so be so much stronger than the effect from the two left arrows.

Secondly, we know that even in the best-performing countries, most students do not receive perfect scores. In PIRLS 2011, for example, the average percentage correct for the highest-performing countries, such as Finland and Hong Kong, is only 68 and 69, respectively (Mullis *et al.*, 2012: 282). Even the highest-performing students in the highest-performing countries do not come close to 100 per cent correct. Thus, in TIMSS 2015 Grade 4 mathematics, in the best-performing countries such as Hong Kong and Singapore, only about 4 per cent of students score 100 per cent correct – contrast this to the 50 per cent or so that are in the highest proficiency category in the same countries (unpublished tabulation, IEA; D. Hastedt, personal communication, March 2017).

Figure 3.5 Counterfactual example from PISA 2015



Source: OECD, 2016.

Conclusions

This paper has presented strong evidence that, when going from lowest levels of overall performance to middle levels of performance, countries make more progress by drastically reducing the proportion of their students that are at the lowest levels of proficiency. They do not, as a rule, increase the proportion of students at high levels of performance by nearly as much. This is demonstrated through simulations using data from several international assessments described above.

In this chapter, we have tried to demonstrate that the pattern of cognitive achievement in low-performing countries is strongly pyramidal, and it is pyramidal both within and between countries. That is, countries' learning achievement distributions have a wide base where the least-performing proficiency level is the most populated, the next higher one is slightly less populated, and so on. The analysis shows, also, that the way countries progress is by reverting the pyramid, by making the pyramid less pyramidal. In fact, in the cases of highest achievement, the pyramid is inverted, with very few students performing at the lowest levels.

In sum, there is a pyramid, and, in moving from low overall performance to middle levels of overall performance, countries do indeed make the pyramid less pyramidal. The results hold across a large variety of assessments in three subjects. The countries in the analysis, however, are not the poorest in the world. It is likely that, were one able to include such countries in the analysis, the bottom of the pyramid would seem even larger, and the path upwards from the bottom would even more emphatically require making the pyramid less pyramidal.

References

Mullis, I.V.S.; Martin, M.O.; Foy, P.; Drucker, K.T. 2012. *PIRLS 2011 international results in reading.* Chestnut Hill, Mass.: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S.; Martin, M.; Loveless, T. 2016. *20 years of TIMSS: international trends in mathematics and science achievement, curriculum, and instruction.* Chestnut Hill, Mass.: TIMSS & PIRLS International Study Center, Boston College, and International Association for the Evaluation of Educational Achievement (IEA).

OECD (Organisation for Economic Co-operation and Development). 2016. *PISA 2015 results (Vol. I): excellence and equity in education.* Annex B1, Ch. 4, Table I.4.1a. Paris: OECD Publishing.

Prahalad, C.K. 2005. *The fortune at the bottom of the pyramid.* Philadelphia, Pa.: Wharton School Publishing.

Pritchett, L. 2013. 'Extreme poverty is too extreme'. Center for Global Development blog, 7 October. Accessed 21 June 2017: www.cgdev.org/blog/extreme-poverty-too-extreme

Simanis, E. 2012. 'Reality check at the bottom of the pyramid'. In: *Harvard Business Review*, June. Accessed 21 June 2017: https://hbr.org/2012/06/reality-check-at-the-bottom-of-the-pyramid

United Nations. 2016. 'Report of the inter-agency and expert group on sustainable development goal indicators'. Statistical Commission Forty-seventh Session, 8–11 March 2016. Retrieved from: https://unstats.un.org/unsd/statcom/47th-session/documents/2016-2-IAEG-SDGs-Rev1-E.pdf

Van Damme, D. 2014. 'How closely is the distribution of skills related to countries' overall level of social inequality and economic prosperity?' ECD Education Working Papers no. 105. Paris: OECD Publishing.

——. 2017. 'How can inter-governmental agencies frame better policy conditions to achieve better "learning at the bottom of the pyramid"?' Paper presented at the Penn Invitational Working Conference on Learning at the Bottom of the Pyramid, Philadelphia, Pa., 3 March.

Wagner, D.A. 2011. *Smaller, quicker, cheaper: improving learning assessments in developing countries.* Paris and Washington: UNESCO IIEP and EFA Fast Track Initiative of Global Partnership for Education.

Wagner, D.A.; Castillo, N.M. 2014. 'Learning at the bottom of the pyramid: constraints, comparability and policy in developing countries'. In: *Prospects* (44), 627–638.

Commentary

Manuel Cardoso

Crouch's paper is a thought-provoking contribution. However, some debatable points should be addressed. First of all, to paraphrase René Magritte…

Ceci n'est pas (forcement) une pyramide…

Crouch asks, 'Is the cognitive distribution pyramidal?' Yes and no. The academic achievement distribution that Crouch examines is pyramidal. But it focuses on results from developing countries or emerging economies in assessments such as TIMSS, PISA, and PIRLS, initially designed for developed countries. For various reasons beyond the space available here, Crouch might answer this question differently if he examined either the distribution *for developed countries* in PIRLS, PISA, or TIMSS, or more importantly for our purposes, the distribution for developing countries *in regional assessments*.

That said, we cannot be literal about the pyramid. The imperative to focus on learning at the bottom of the pyramid is rights-based and linked to a commitment to the most marginalized. From that perspective, this concern is legitimate regardless of the number of children in the bottom categories. In this sense, Crouch introduces a useful distinction between absolute and relative academic poverty.

Some children are learning less, academically, than others; this is a problem because it reflects unequal opportunities to learn. This is *relative* academic poverty. But some children see their academic chances truncated at an early age because they have not mastered the foundational skills that allow them to continue learning in school. This is an even bigger problem; this is *absolute* academic poverty. This not only puts them at a disadvantage in relation to others, it also does not allow them to survive in school. All children should attain basic reading and numeracy skills that allow them to continue learning, not only as children, but throughout their lives. This brings me to the next contentious point. To paraphrase J.R.R. Tolkien…

One to rule them all.

Crouch criticizes the tendency to focus on a single indicator such as 'the average test score …, or the percentage of students above a certain threshold' This conflates two issues. The first issue is: should we have

one indicator or more than one? It depends on purpose. Advocacy may prioritize a lead indicator, while technical analysis may require a battery.

The second issue, also relevant, is: if we have only one indicator, or even if we have many but prioritize one, which one should it be? Crouch argues that focusing on one indicator prevents us from taking equity into account. This is true of 'the average test score'. But it is not necessarily true of 'the percentage of students above a certain threshold'. It depends on where that threshold is located. A threshold that is low enough to distinguish between academic haves and have-nots can highlight equity issues. It can show the need to reallocate resources and devise targeted strategies.

UNICEF has just developed such an indicator for reading, and another one for mathematics, both of which focus on the percentage of *children*, not just students, that have developed the foundational learning skills required to continue to move forward in school. A few countries are already piloting data collection for this indicator through the Multiple Indicators Cluster Survey (MICS). Being household-based, MICS reaches children both in and out of school; it is also administered in many countries that do not participate in cross-national assessments (UNICEF, 2017). UNICEF's Foundational Learning Skills module also has the potential, in principle, to be adapted to other household-based platforms.

In summary, whether the pyramid is literal or not should not be our main concern. Our priority should be reaching all children in all countries, regardless of whether they are in school or not. We should reach them with both assessments that will help us identify gaps and raise awareness, and formal and non-formal education strategies to address their needs.

References

UNICEF. 2017. 'Collecting data on foundational learning skills and parental involvement in education'. MICS Methodological Paper no. 5. Retrieved from: http://mics.unicef.org/publications/methodological-papers

Commentary

Emily Hannum

This chapter considered some of the major initiatives to assess literacy and learning around the world. It argued, with evidence from cross-national survey data, that countries go from the lowest levels of overall performance to middle levels of performance by reducing the proportion of students that are at the lowest levels of proficiency.

Reflecting on the conference and the findings presented in this chapter provoked some questions about how data sources handle complex questions of language. What have been the protocols for adapting common literacy assessment items into different languages, and how are equivalencies addressed? I imagine that this question is one that test designers have grappled with extensively. It would be useful to hear a bit more in the paper about this issue, and its implications for interpreting findings. Even more complicated is the issue of whether and to what degree language issues might play a part in whether children land at the 'bottom of the pyramid' in literacy assessments. Children for whom school language is not home language may be disadvantaged in assessments given in the national language, but may also be disadvantaged in assessments given in their mother tongue, if their formal learning has all taken place in the national language. It may not even be possible to take an assessment in one's home language if the home language is largely used for speaking and is not written.

Further, in multilingual national contexts, some research suggests that literacy in the national language could bring both symbolic and instrumental benefits that differ substantially from those associated with literacy in a minority language. In short, it seems important in studies of this kind to elaborate on how assessments have handled questions of equivalencies across language, the degree to which there is agreement across these assessments about the language in which literacy is being assessed, and protocols for dealing with minority languages.

Finally, the conference discussion raised the idea to study countries that have reduced or eliminated the BoP to see what can be learned from their assessment practices. An example given was the Republic of Korea. This leads one to wonder how to tease out whether the 'elimination' of the BoP can be linked to assessments, to features of

the school system, or to features outside the school system altogether, such as changes in family conditions associated with dramatic economic development, rapid increases in the educational composition of parents, and the rising competition for educational credentials that has driven high levels of participation in extra-curricular tutoring. While the Republic of Korea may represent an extreme case, it raises the broader issue of how to account for the possibly crucial importance of factors outside the purview of assessment, and even outside the purview of the educational system, in driving reductions at the BoP.

Chapter 4
Use of international assessments to support learning at the bottom of the pyramid

Anil Kanjee

Introduction

The last two decades have seen a significant increase in the number of large-scale international studies implemented across the globe, as well as the number of countries participating in these studies. While reasons for this vary widely, a key concern raised by a number of scholars pertains to the limited information available on the impact of these studies on education systems across different countries (Addey *et al.*, 2017; Alexander, 2015; Gorur, 2017). While these studies have the potential to significantly impact what and how we teach (Abu-Alhija, 2007; Alexander, 2015; Schiefelbein and Schiefelbein, 2003), there has been limited focus on the impact of these studies on the learning and teaching process in general (Baird *et al.*, 2016), and in particular, its impact on poor and marginalized learners and/or countries at the bottom of the pyramid (Wagner and Castillo, 2014).

To contribute to current debates on the value and use of international assessment studies (IAS), this chapter aims to explore the extent to which IAS support learning at the bottom of the pyramid (BoP). The chapter begins by reviewing the purpose of IAS and the reasons that countries participate in these studies. Next, an overview of IAS's usefulness in enhancing learning and teaching is explored, followed by a review of the likely impact of these studies on supporting BoP learning achievement. The chapter concludes by listing possible alternatives that need to be explored for addressing learning needs of poor and marginalized learners.

Purpose of IAS and reasons for participating

Across the IAS reviewed, one of the key reasons noted for conducting IAS is to facilitate the sharing of information, knowledge, and experiences that may be used by countries to improve their education systems (Howie, 2012; Lockheed, 2012; Reddy, 2005). This is a reasonable and noteworthy exercise given the possibilities of obtaining new ideas, and relevant evidence, on how best to identify and address

specific policy challenges for improving learning within the education system. The Latin American Laboratory for Education Evaluation (LLECE) specifies that its work is a 'key instrument to monitor and follow up the Education 2030 Agenda Frame of Action and Sustainable Development Goal 4' (UNESCO Santiago, 2017). TIMSS and PIRLS (2017) note that they 'enable participating countries to make evidence-based decisions for improving educational policy', and SACMEQ (2017) states that participation can 'generate information that can be used by decision-makers to plan the quality of education'. Moreover, many of these organizations believe that a key objective is developing the capacity of participating countries to conduct, and use effectively, these evaluations (Lockheed and Wagemaker, 2013).

However, in her review of what drives participation in IAS, Addey (2015) notes that reasons for participation go well beyond accountability and policy. For example, countries may participate:

- to meet requirements of funders and/or to establish systems of accountability for their educational systems;
- to compare themselves to other countries, so as to evaluate how far they need to go to catch up in terms of skills;
- to obtain a form of legitimacy and credibility, by being part of a group of countries that value public education;
- for the 'prestige' of competing and benchmarking themselves against the exclusive club of rich countries.

The key point is that specific reasons for participation in IAS, as well as the primary purpose, underlying philosophy, and specific approach of different international agencies responsible for implementing these studies, determine the extent to which IAS can help countries improve learning and teaching, and in particular, address the learning needs at the BoP.

Value of IAS for improving learning and teaching

A review of the literature regarding the value and use of IAS for supporting improvements in learning indicates that the primary value, particularly with regard to learning at the BoP, lies in providing information for use in policy change. In practice, this implies that these studies can only have an indirect, long-term impact on improving learner achievement. However, whether and how this 'benefit' translates to practice depends on a range of factors and varies substantially between and within countries, as

well as across different IAS. These factors are discussed below under the following headings: comparability, learning and teaching, policy/practical, technical and data, financial/human resources, and promoting performativity.

Challenge of comparability

Beaton *et al.* (1999) note that through conducting comparisons across different education systems, IAS provide countries with valuable information for identifying key factors that impact learning achievement. However, the authors also caution against 'simplistic and/or selective interpretations that fail to take account of the full picture – especially the complex and multi-factorial nature of the educational process' (Beaton *et al.*, 1999: 15–16). Similarly, Alexander (2010) describes a number of difficulties in cross-national comparisons when identifying explanations for countries performing well or poorly, and argues for the need to 'dig deeper' and explore a range of factors that extend beyond the education system, including cultural and linguistic homogeneity, and low rates of immigration.

Highlighting Alexander's point, Carnoy (2015: 3) provides the following exemplar:

> The OECD has repeatedly held up Shanghai students and the Shanghai educational system as a model for the rest of the world and as representative of China, yet the sample is not representative even of the Shanghai 15-year-old population and certainly not of China. In addition, Shanghai schools systematically exclude migrant youth. These issues should have kept Shanghai scores out of any OECD comparison group.

Carnoy and Rothstein (2015) also question the validity of any comparisons between vastly different countries. The authors highlight the significant differences between the USA and high-performing countries and areas like Finland, Singapore, and Shanghai with regards to factors such as population size, the structure of the education system and its implementation, cultural diversity, variations in education funding and provision, and percentage of population represented within the different socio-economic subgroups in the country. Similarly, Baird *et al.* (2016) and Wagner (2010) note that the challenges of sampling learners from under-represented subgroups (e.g. disabled learners, children of migrants or from poor families, a lack of correspondence between the school curricula of participating countries, effects of motivation, or issues of item translation) all tend to make comparisons between countries problematic.

Learning and teaching challenges

A key rationale noted by countries participating in IAS is to '*identify lessons learnt for improving learning and teaching*' (Addey *et al.*, 2017; Alexander, 2012). This, however, assumes that IAS provide relevant and useful information on learning and teaching, an assumption that is not necessarily valid. In his review of IAS, Wiliam (2008) argues that the tests used in these studies are generally not sensitive to instruction, and thus provide limited information on student learning. Wiliam (2008) lists three key reasons that IAS are insensitive to instruction:

1. 'learning itself is relatively insensitive to instruction in that the progress made by individual students is rather slow compared to the variability within the cohort' (p. 254);
2. 'the procedures used in almost all test construction decrease the sensitivity of the test to instruction' (p. 255);
3. 'the specific procedures used to develop tests for international comparisons decrease the instructional sensitivity yet further' (p. 255).

In his review of PISA and TIMSS, Morris (2016) argues that these studies can provide some useful information to countries, and also provide governments with a powerful source of legitimacy for education reform. However, with respect to supporting learning improvements, Morris notes that 'the quest for improved PISA/TIMSS scores has resulted in massive reform programmes that often have had no impact on pupil learning outcomes' (2016: 6). Similarly, in her review of the role of international large-scale assessments in developing countries, Lockheed (2012) also argues that IAS are poorly suited for improving the individual performance of students.

In their review of the link between assessment and learning, Baird *et al.* (2016) contend that while IAS have the capacity to inform us regarding learning, few researchers have used the results in that way. The authors highlight four barriers that hinder researchers' use of data from IAS for investigating students' understanding, and thereby prevent them from contributing to the knowledge base of their learning:

1. Keeping test items confidential – which 'makes it difficult to make sense of what students have learned if you cannot see what they have been asked to do' (p. 26).
2. Plausible values are not transparent, meaning that 'without being able to trace back the connections between plausible values,

raw scores and the content of items, concrete conclusions about learning cannot be drawn' (p. 26).

3. The availability of better data from national datasets due to the restricted range of topics covered in IAS, or lack of data available to address the specific questions of researchers. However, the authors note that this is not necessarily true of all IAS.
4. Comparisons between countries may be problematic given the range of factors that affect learning (see previous section). This, the authors argue, means that 'using the data to compare countries to draw conclusions about how well students are learning in different education systems is not straightforward' (p. 27).

While Baird *et al.* (2016) argue for more research and greater scholarship to explicate the relationship between assessment and learning, they also acknowledge that IAS 'have not yet taught us much about learning or contributed to theories of learning, but they have had an impact upon what is learned and how it is learned through education policy' (Baird *et al.,* 2016: 29).

Policy/practical challenge

In their summary of a seminar series on '*The potentials, politics and practices of international educational assessment*', Maddox and Addey (2016: 2; see also this volume, *Chapter 4*) note that '"good" assessment should support informed policy making that is sensitive to the characteristics of diverse societies, and understands these characteristics as complex and multi-dimensional. It should support informed and equitable and nuanced policy making processes and not create simplistic understandings or unwanted or ill-informed policy shocks'. In her review of the role of IAS in developing nations, Lockheed (2012) argues that the value of these studies lies in motivating relevant policy reforms that can impact teaching and learning, and creating learning environments for local experts to improve their technical skills as well as national assessment systems and practices. This rationale for participating and/or implementing IAS – that is, for obtaining evidence to inform policy – is widely accepted among researchers and policy-makers (Addey *et al.*, 2017; Alexander, 2012; Howie, 2012; Lockheed and Wagemaker, 2013). However, Addey *et al.* (2017) contend that national, economic, and political contexts within countries greatly impact how data from IAS are used, and thus IAS are often used to support specific government agendas that could include specific reforms or even inaction.

However, in practice the availability of valid, reliable, and relevant data does not necessarily mean that the data will be used effectively to inform the policy process. Commenting on how policy-makers analyse and interpret information from IAS, and how they translate what they discover into practice, Alexander (2012: 4) highlights three tendencies: '(i) extreme selectivity in the evidence cited; (ii) a misplaced faith in the ability of national interventions to transform classroom practice and hence raise standards; (iii) a preference for interventions that are high stakes and draconian'. Similarly, in their review on the use of data from IAS in the USA, Carnoy and Rothstein (2015: 124) note that policy-makers not only 'typically ignore the social and economic causes of low achievement, but international test comparisons based on average national scores or current achievement gaps encourage policy-makers to draw misguided conclusions about school improvement and may actually lead them to pursue inappropriate and even harmful reforms'.

Technical and data challenges

While IAS are renowned for adhering to extremely high standards, these studies are still prone to a number of technical and data challenges. The estimation model used to analyse results as well as the population coverage of samples selected are two key issues, highlighted by Rutkowski and Rutkowski (2016), that can impact the validity and reliability of the data for making comparisons across countries. In their review of PISA results, the authors report that the fundamental assumption of item parameter equivalence, which allows for meaningful cross-cultural comparisons, does not hold. They note that the consequence of this measurement error 'is that achievement rankings may not be accurate and system-level comparisons can lead to incorrect conclusions regarding achievement differences' (Rutkowski and Rutkowski, 2016: 254). In their analysis of the TIMSS 2003 data for South Africa, Gregory and Kanjee (2009) also report a preponderance of item misfit, indicating a lack of scale and measurement equivalence, and further argue that the country's science achievement results should not be compared to other countries using the TIMSS metric. Rutkowski and Rutkowski (2016) also note that inadequate coverage of the population leads to sampling errors that not only impact the generalizability of the results, but also the inferences that can be made from the data. Highlighting the differential representation of students from socio-economic classes in countries in the PISA study, Carnoy (2015) argues that making inferences about the quality of educational systems via the ranking of countries by average test score is misleading in this context. In addition, a key issue raised by Goldstein

(1995, 2004) pertains to the challenges of creating unidimensional tests that can accurately summarize different aspects of achievement into a single score. In this context, Goldstein questions both the validity of the learner performance scores as well and their accuracy when conducting comparisons across different socio-economic and cultural contexts.

Financial and human resources

Another key factor for participating in IAS for many developing nations pertains to the costs of these studies, which include the payment of fees, implementation costs, as well as costs related to planning such studies and producing reports and results. While information on costs of participation is not readily available and cannot be easily estimated, Wagner *et al.* (2011: 10) note that in times of fiscal constraint, 'Learning about education has to be balanced against *what is learned, for what purposes, and at what cost*. The evaluation of assessment costs is an issue that will need considerably greater attention in the field of international education.'

A key requirement for participating in IAS, which is especially significant and taxing for low- and middle-income countries, pertains to the human resources required to ensure effective implementation of these studies. Often, it is only a small group of experts, in either the education ministry or associated research institutions or universities, who are involved in these studies. Demands on these experts are often excessive, especially when countries participate in multiple studies (e.g. Zambia – SACMEQ and PISA for development, South Africa – TIMSS, PIRLS, SACMEQ), and when relevant staff are also responsible for other projects. In these instances, such staff have limited time for detailed analyses, reporting, and dissemination of the most recent data, and are often required to begin preparation for the next round of the study before data from the current study can be properly analysed or 'used'. To some extent, these phenomena also help to explain the limited publications and secondary analysis available from many poor and middle-income countries participating in IAS.[10]

Promoting performativity

The predominant use of measurable criteria to monitor and evaluate the functioning of an education system and the performance of key

10. It should be noted that in South Africa, the coordinators of TIMSS have recently expanded efforts for ensuring greater use of the data and involvement of larger stakeholders in the project.

role-players within that system has been gaining momentum over the last two decades, with more and more countries adopting this approach. Within such measurement-driven systems, also referred to as *performativity* based systems, test scores from large-scale assessments, examinations, and standardized assessments play a critical role in providing the 'evidence' to determine how well the system is performing. Performativity has been defined 'as a technology, a culture, and a mode of regulation that employs judgments, comparisons and displays as means of incentive, control, attrition and change—based on rewards and sanctions (both material and symbolic)' (Ball, 2003: 2016). Ball further argues that performativity is the translation of complex social processes and events into simple figures or categories of judgment.

Gorur's (2016) review of the impact of PISA on education systems demonstrates how cultures of performativity support the development of assessment systems that are characterized by 'a reliance on numbers; enhanced mechanisms of accountability; a heightened focus on education as an economic commodity; a proliferation of testing; and the viewing of education as a global race with winners and losers' (Gorur, 2016: 608). Two key practical consequences of this approach are highlighted by Gorur: (a) it promotes a new standardization that facilitates the 'interpretation of complex issues into simply stated problems that suggest simplistic solutions which are measureable and can be tracked over time' (Gorur, 2016: 609); and (b) the focus on a few core subjects for measuring the education system, i.e. literacy and numeracy, while allowing for easier monitoring and evaluation of teachers and schools, also places emphasis on comparison and competition.

In the context of education in South Africa, similar performativity regimes, or what Jansen (2001) calls 'politics of performance', and what Kanjee and Sayed (2013) refer to as a 'measurement driven assessment system', have manifested in the signing of delivery agreements between the Department for Performance Monitoring and Evaluation, located in the office of the presidency, and all government ministers. In the case of education, this agreement formed the basis of the *Action Plan to 2014: Towards the realisation of Schooling 2025* (South Africa, 2012). This plan lists five key goals and 14 indicators associated with clear targets and timeframes. Of the 14 targets noted, the data source for six is the South African Annual National Assessments,[11] for four is the National Senior

11. These assessments were subsequently stopped in 2015 due to pressure from teacher unions.

Certification Examinations, for three is international assessment studies and for one is teacher assessments – targets and information for which are yet to be specified. For Jansen (2005) such regimes – preoccupied with outcomes and results – are also intended to absolve the educational departments of their key responsibility of ensuring access to quality education for all. The real victims of such regimes, according to Jansen, are schools and communities comprising poor and marginalized learners – that is, those at the BoP, who already have to endure extremely challenging contexts that negatively impact their chances of obtaining good-quality education.

Likely impact on the BoP

Given the challenges highlighted, it is highly unlikely that IAS can have any major impact on addressing the specified learning needs of poor and marginalized learners and/or countries – those at the BoP. Even in instances when IAS provide information that accounts for social class composition, or that allows for more nuanced analysis regarding poor and marginalized learners (Carnoy and Rothstein, 2015), there is still the challenge of effective use of information, and its interpretation for use in policy and practice. A key challenge noted by Wagner and Castillo (2014) is the limited information available regarding learning at the BoP. Given the domination of developed nations in almost every aspect related to IAS, the authors note that this situation is unlikely to change much in the future. Specifically, Wagner and Castillo (2014: 633) state that 'if comparability is the central goal in IAS, less attention maybe be paid to the local and cultural validity of the definitions and classifications of learning… and that data may become less meaningful and potentially less applicable at the local level'. The authors pose a key question – 'Can international goals and commensurate statistics be maintained as stable and reliable if localized approaches are chosen over international comparability?' (p. 634) – and allude to an 'answer' when they note that 'Which types of comparability are most important depends on the policy goals desired, as well as timing and resource considerations' (p. 634), which, in the current context of IAS, reside in the hands of developed countries.

Alternative assessment strategies to consider

Given the ubiquitous nature of national assessments and IAS, as well as their growing influence (and impact) on education systems, especially with respect to poor and marginalized learners, a key question to address

is how to better support learners at the BoP. Specifically, the key dilemma is whether to 'organize', 'advocate', and/or 'wait' for relevant systemic reform, or to promote immediate steps that can minimize, at least in the short term, some negative effects of current performativity regimes, or to do both. In their proposal for a possible solution, Wagner and Castillo (2014: 634) argue for 'hybrid assessments that offer localized comparability that large-scale assessments do not, and can offer more focused results for improving learning and interventions among poor and disadvantaged populations'. In practice, this call for assessments that function alongside, or instead of, current performativity regimes can also directly impact learning and teaching processes, especially those that affect poor and marginalized learners. Within this context, the application of an 'Assessment for Learning' approach merits further consideration.

Assessment for Learning (AfL) is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go, and how best to get there (Assessment Reform Group, 2002). Baird *et al.* (2016) note that the theory and philosophy of the AfL approach motivates practices that, at their root, impact learning. Thus, such approaches might have a direct impact on daily teaching and learning practices within the classroom, unlike IAS, where the connection between the assessments and learning is less direct because it is mediated through policy, curriculum, and assessment design.

An underlying assumption of the AfL approach is the formative use of assessment information, defined by Wiliam (2010: 43) as 'the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have made in the absence of that evidence'. In practice, effective implementation of the AfL approach requires two key processes: (a) implementation of formative assessment during the teaching and learning process; and (b) the formative use of summative data for improving learning and teaching.

Formative assessment during the teaching and learning process

Research by Wiliam (2010) indicates that the effective use of formative assessment strategies in the classroom results in not only increased learning gains, but gains that are observed among all children, including those from poor and marginalized backgrounds. As a pedagogical approach to improve learning, AfL focuses on strategies and techniques

that ensure all learners are clear about the learning intentions and success criteria, and that teachers use relevant tools and techniques to enhance engagement to include all learners in the classroom, especially those that do not participate, and to identify – and, if possible, address – learning gaps (Wiliam and Thompson, 2007). In addition, the effective use of oral and written feedback can help learners address specific learning gaps, while the use of peer- and self-assessment strategies can empower learners to take responsibility for their own learning (Wiliam, 2010).

Notwithstanding its potential to address the learning challenges for poor and marginalized children, a key challenge with the use of AfL strategies pertains to the limited information on how to scale up such programmes, given the high levels of instructor knowledge and capacity required to implement AfL effectively in the classroom. In part, this is due to the limited research on the use of AfL at scale, but it also reflectsthe lack of these studies from developing nations, specifically countries at the BoP. However, recent findings from a large-scale randomized control trial to determine the impact of an AfL professional development programme on teaching and learning in South African schools show positive results. The programme was implemented in a random sample of schools drawn from all poverty quintiles, where lower quintiles (i.e. Quintile 1) represent poorly resourced and generally low-performing schools with mainly underqualified teachers and learners from low socio-economic backgrounds (i.e. learners at the BoP). Higher quintiles (i.e. Quintile 5) represent well-resourced and generally higher or better-performing schools with better-qualified teachers, and learners from middle to high socio-economic backgrounds. Results from the mid-term evaluation revealed significant improvements in AfL knowledge among teachers across all poverty quintile schools, while evidence of enhanced learner engagement across all school types in the treatment groups was also noted[12] (Kanjee, 2016). While the impact on learner performance will only be determined at the end of the project, what these findings ultimately point to is that in order for the specific learning needs of poor and marginalized children to be addressed, solutions need to be sought in classroom-based practices, and not in any form of IAS.

Formative use of summative results

Within the context of South Africa, Kanjee and Moloi (2016) explored the feasibility of using a standards-based approach for reporting

12. Year 2 of the study was expected to be completed at the end of 2017.

assessment results from a range of different sources, including IAS, national and provincial assessments, school-based tests and standardized examinations. In practice, this allows for:

- greater use of data from standardized assessments that currently comprise the dominant source of data within many education systems;
- reporting results to highlight key learning needs of all learners, including those at the BoP;
- providing relevant information that helps teachers, school leaders, and education officials identify and set specific targets that focus on improving learning (and teaching) in schools and communities with large proportions of poor and marginalized children.

First, we argue that, contrary to current practice where learner performance is reported using a single number (usually a percentage score), the use of a standards-based approach provides detailed information on what learners functioning at different performance levels know and can do. Using this information, teachers, parents, school leaders, district officials, and policy-makers can develop relevant interventions that support the specific learning needs of all learners, focusing specifically on those at the lower levels of performance. Second, we suggest that providing information on learners functioning at the different levels of performance allows for monitoring as well as for setting targets – at the classroom, school, and district levels – that focus specifically on reducing the percentage of learners functioning at the lowest levels of performance.

Even more important, however, this approach can combat the current negative consequences of performativity regimes within the South African education system, where a large number of low-performing learners are excluded from examinations; regular schooling is suspended to prepare for national assessment; and where schools and teachers prioritize support to ensure higher-performing learners improve their results, often at the expense of poor and marginalized learners. However, we caution that for such an approach to work, adequate support must be provided to teachers and schools, and any monitoring that is conducted must develop and promote a culture of improving learning for all as opposed to improving performance. In practice, this will require a significant change in how ministry officials operate – from one where the key questions during any school visit are: 'How much of the curriculum have you completed to date?', 'Are all your files up to date?' and, 'How much has the average score for your school

improved?' to: 'How can I assist to improve what is working and address what is not working in your school or classroom?' and 'How many learners at the lowest performance levels have you assisted to improve their performance?'

Conclusion

While the value and use of IAS has been an area for research and discussion for several decades now, there has been limited focus on the impact of these studies on the learning and teaching process in general, and in particular, its impact on the learning and teaching of poor and marginalized learners – those at the BoP. This is despite the significant increase in the number of international (as well as national) assessment studies that have been conducted over the last two decades. A review of the literature indicates that IAS provided limited opportunities for improving learning in general, and learning at the BoP in particular. Key reasons for this include lack of valid data on learning and teaching, the limited availability of resources, and the increasing dominance of performativity regimes within countries.

Notwithstanding the enormous challenges countries face in education, specific solutions to address learning needs at the BoP lie in the use of 'Assessment for Learning' approaches. However, these approaches will require massive investment by the stakeholders who matter the most, and who have the greatest impact on learning in schools at the BoP – the teachers, school leaders, parents, and learners themselves.

References

Abu-Alhija, F.N. 2007. 'Large-scale testing: benefits and pitfalls'. In: *Studies in Educational Evaluation*, 33, 50–68.

Addey, C. 2015. *What drives participation in international educational assessments?* Laboratory of International Assessment Studies. Retrieved from: http://international-assessments.org/what-drives-participation-in-international-educational-assessments/

Addey, C.; Sellar, S.; Steiner-Khamsi, G.; Lingard, B.; Verger, A. 2017. 'The rise of international large-scale assessments and rationales for participation'. In: *Compare*, *47*(3), 434–452.

Alexander, R.J. 2010. 'World class schools – noble aspiration or globalised hokum?' In: *Compare*, *40*(6), 801–817.

——. 2012. 'Moral panic, miracle cures and educational policy: what can we really learn from international comparison?' In: *Scottish Educational Review, 44*(1), 4–21.

——. 2015. 'Teaching and learning for all? The quality imperative revisited'. In: *International Journal of Educational Development*, 40, 250–258.

Assessment Reform Group. 2002. *Assessment for Learning: 10 principles*. Cambridge, UK: University of Cambridge Assessment Reform Group.

Baird, J.; Andrich, D.; Hopfenbeck T.; Stobart, G. 2016. 'Assessment and learning: fields apart'. Unpublished manuscript, Oxford University Centre for Education Assessment.

Ball, S.J. 2003. 'The teacher's soul and the terrors of performativity'. In: *Journal of Education Policy*, *18*(2), 215–228.

Beaton, A.E.; Postlethwaite, T.N.; Ross, K.N.; Spearritt, D.; Wolf, R.M. 1999. *The benefits and limitations of international educational achievement studies.* Paris: IIEP/UNESCO.

Carnoy, M. 2015. *International test score comparisons and educational policy: a review of the critiques*. Boulder, Colo.: National Education Policy Center. Retrieved from: http://nepc.colorado.edu/ publication/ international-test-scores

Carnoy, M.; Rothstein, R. 2015. 'What international test scores tell us'. In: *Society*, *52*(2), 122–128.

Goldstein, H. 1995. 'Interpreting international comparisons of student achievement', report no. 63. Paris: UNESCO.

——. 2004. 'International comparisons of student attainment: some issues arising from the PISA study'. In: *Assessment in Education*, *11*(3), 319–330.

Gorur, R. 2016. 'The potentials, politics and practices of international educational assessment', discussion paper for Laboratory/ESRC Seminar, Humboldt University, Berlin, 15–16 September.

——. 2017. 'Towards productive critique of large-scale comparisons in education'. In: *Critical Studies in Education*, *58*(2), 1–15.

Gregory, K.; Kanjee, A. 2009. 'Modelling science achievement in South Africa using international achievement data: some comparability and validity issues', paper for 35th annual meeting of International Association for Educational Assessment, Brisbane, Australia. 13–19 September.

Howie, S. 2012. 'High stakes testing in South Africa: friend or foe?' In: *Assessment in Education, 19*(1), 81–98.

Jansen, J.D. 2001. 'On the politics of performance in South Africa'. In: *Prospectus, 31*(1), 553–562.

——. 2005. 'Targeting education: the politics of performance and the prospects of "Education for All"'. In: *International Journal of Educational Development, 25*(4), 368–380.

Kanjee, A. 2016. 'Preliminary findings of the AfL RCT capacity development programme'. Unpublished report, Tshwane University of Technology, Pretoria, South Africa.

Kanjee, A.; Moloi, Q. 2016. 'A standards-based approach for reporting assessment results in South Africa'. In: *Perspectives in Education, 34*(4): 29–51.

Kanjee, A.; Sayed, Y. 2013. 'Assessment policy in post-apartheid South Africa: challenges for improving education quality and learning'. In: *Assessment in Education, 20*(4), 442–469.

Lockheed, M.E. 2012. 'Policies, performance and panaceas: the role of international large-scale assessments in developing countries, Compare forum'. In: *Compare, 42*(3), 512–518.

Lockheed, M.E.; Wagemaker, H. 2013. 'International large-scale assessments: thermometers, whips or useful policy tools?' In: *Research in Comparative and International Education, 8*(3), 296–306.

Maddox, B.; Addey, C. 2016. 'The futures and promises of international educational assessment'. Seminar series report on 'The potentials, politics and practices of international educational assessment', Laboratory of International Assessment Studies. Retrieved from: http://international-assessments.org/what-drives-participation-in-international-educational-assessments/

Morris, P. 2016. *Education policy, cross national tests of pupil achievement, and the pursuit of world-class schooling: a critical analysis.* London: UCL Institute of Education Press.

Reddy, V. 2005. 'Cross-national achievement studies: learning from South Africa's participation in the Trends in International Mathematics and Science Study (TIMSS)'. In: *Compare, 35*(1), 63–77.

Rutkowski, L.; Rutkowski, D. 2016. 'A call for a more measured approach to reporting and interpreting PISA results'. In: *Educational Researcher, 45*(4), 252–257.

SACMEQ. 2017. 'The Southern and Eastern Africa Consortium for Monitoring Educational Quality mission'. Retrieved from: www.sacmeq.org/?q=mission

Schiefelbein, E.; Schiefelbein, P. 2003. 'From screening to improving quality: the case of Latin America'. In: *Assessment in Education, 10*(2), 141–154.

South Africa. 2012. Department of Basic Education. *Action Plan to 2014: towards the realisation of Schooling 2025 – full version*. Pretoria: Department of Basic Education.

TIMSS; PIRLS. 2017. 'About TIMSS and PIRLS International Study Centre'. Retrieved from: https://timssandpirls.bc.edu/about.html

UNESCO Santiago. 2017. 'What is the Latin American Laboratory for Assessment of the Quality of Education (LLECE)?' UNESCO Regional Education Office for Latin America and the Caribbean. Retrieved from: www.unesco.org/new/en/santiago/education/education-assessment-llece/

Wagner, D.A. 2010. 'Quality of education, comparability, and assessment choice in developing countries'. In: *Compare, 40*(6), 741–760.

Wagner, D.A.; Babson, A.; Murphy, K.M. 2011. 'How much is learning measurement worth? Assessment costs in low-income countries'. In: *Current Issues in Comparative Education, 14*(1), 3–21.

Wagner, D.A.; Castillo, N.M. 2014. 'Learning at the bottom of the pyramid: constraints, comparability and policy in developing countries'. In: *Prospects,* 44, 627–638.

Wiliam, D. 2008. 'International comparisons and sensitivity to instruction'. In: *Assessment in Education, 15*(3), 253–257.

——. 2010. *Embedded formative assessment*. Bloomington, Ind.: Solution Tree Press.

Wiliam, D.; Thompson, M. 2007. 'Integrating assessment with learning: what will it take to make it work?' In: D. Dwyer (Ed.), *The future of assessment: shaping teaching and learning* (pp. 53–82). New York: Lawrence Erlbaum Associates.

Commentary

Amita Chudgar

Can international assessment studies (IAS) support learning at the BoP? According to Anil Kanjee, these studies may have an 'indirect, long-term impact' but their ability to inform day-to-day practices is limited. Kanjee lists several factors that lead him to this conclusion. In particular, two limitations of IAS stand out. First, IAS often fail to gather nuanced information on different types of marginalization, which limits the ability of such studies to identify and observe all the children at the BoP. Second, IAS generally report aggregated student performance and do not provide item-specific performance information. Such information is essential if we want to understand where the learning gaps are occurring for those at the BoP. These dual limitations – not being able to identify all those at the BoP, or to understand where their learning gaps are – limit the utility of IAS to support learning at the BoP.

To address the specific learning needs of children at the BoP, Kanjee proposes Assessment for Learning (AfL), a classroom-based assessment approach. This approach is designed to be close to the students and attentive to their learning experiences. It therefore overcomes both the challenge of identifying the most marginalized, and of understanding the gaps in their learning. As Kanjee discusses the specifics of AfL, it becomes apparent that teachers and school leaders, especially teachers and school leaders of those at the BoP, are central to its successful execution.

So what do we know about teachers and school leaders who work at the BoP? Our knowledge about the teachers of children who are most marginalized is limited but growing (Luschei and Chudgar, 2016; see also a related discussion in Chudgar, Chandra, and Razzaque, 2014). We understand that teachers who teach those at the BoP are often younger, less experienced, and often receive uneven pre- and in-service training and support. These teachers work in challenging circumstances, with limited resources, often in multi-grade classrooms, and often with several competing demands on their time that may span from operating the school to teaching the children. We know far less about the school leaders in these marginalized circumstances, but it may be reasonable to extrapolate that their profiles are similar.

This profile of teachers (and school leaders) at the BoP makes it hard to envision how any local, bottom-up response (such as AfL) to

the top-down IAS approach can be implemented and scaled effectively. Kanjee calls for 'massive investment by stakeholders that matter the most'. Indeed, not just a massive, but also a thoughtful investment has to be made in teacher training and mentoring to equip them to teach in challenging circumstances, and in teacher recruitment and retention in marginalized communities for the success of approaches like AfL.

The relevance and limitations of IAS at the BoP have many facets, ranging across politics, economics, and psychometrics. A crucial aspect of this discussion must also be the important role that teachers and school leaders play in any assessment efforts at the BoP, and by extension the strengths and limitations of a nation's teacher training, recruitment and retention policies, and the existence or lack thereof of qualified and motivated teachers at the BoP. Among the many contributions that Kanjee makes in his paper, this is one of the key emerging insights.

References

Luschei, T.F.; Chudgar, A. 2016. *Teacher distribution in developing countries: teachers of marginalized students in India, Mexico, and Tanzania.* New York: Palgrave Macmillan.

Chudgar, A.; Chandra, M.; Razzaque, A. 2014. 'Alternative forms of teacher hiring in developing countries and its implications: a review of literature'. In: *Teaching and Teacher Education*, 37, 150–161.

Commentary

Michael G. Fast

International assessments tell us next to nothing about learning achievement in most participating countries (as noted, for example, in Dr Kanjee's reference to China's performance), let alone in those countries at the BoP. To qualify this statement, there are many local prerequisites that must be met for such assessments to even begin to offer some value, which I briefly present here. Anil Kanjee has the better deal here by having to argue for the threats of such assessments, given that from my perspective and clearly his, the threats and challenges are numerous and real.

On 6 May 2014, 83 academics from across the world sent a public letter to the director of the OECD expressing grave concern about the negative consequences of PISA tests.[13] They identified 'rankings' as the major *raison d'être* of international assessments; Dr Kanjee identifies them as 'league tables that rank order achievement scores'. The public letter made a number of key criticisms, and I refer to them here because Kanjee's paper rightly identifies the same range of problems: quick policy fixes that do not lead to sustainable change; a narrow focus on measurable aspects of education; bias towards the economic value of education; impoverished instruction caused by focus on assessments; insidious involvement of public–private partnerships that could stand to gain locally from the results. To sum up, the letter states that 'the new PISA regime, with its continuous cycle of global testing, harms our children and impoverishes our classrooms'. Kanjee's key point is along the same lines: unless countries have the requisite capacity for further analysis and willingness to use this information effectively, it is highly unlikely that these studies can have an impact on any aspect of learning (and teaching) within countries, and less so for learners at the BoP.

This takes me to where I would like to make a contribution to Kanjee's conclusions, as an implementer and a proponent of standardized testing in many of the countries that might be described as 'bottom of the pyramid'. Here are some observations from the MIDEH[14] project in Honduras, a USAID-funded national standards and assessment project that is moving towards its 14th year of implementation:

13. www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics
14. www.air.org/project/mideh-honduras-mejorando-el-impacto-al-desempe-o-estudiantil-de-honduras

1. Top-performing schools in Honduras, based on the National Assessments in math and Spanish, have often been multi-grade rural schools.
2. District mayors in targeted areas of Honduras use educational data from National Assessments to set school and district targets to guide instructional improvement.
3. In 2004, the grass-roots educational profession in Honduras manifested itself against standards and against external assessment, particularly the unions; in 2017, the educational profession across the country owns and implements (and has done for 7 years or so) national standards aligned with the curriculum, as well as the formative and summative assessment system that allows them to find out how students are doing and how they can set targets for improvement.
4. End of the year assessments are administered in November and the results are published in customized reports targeting teachers, schools, district authorities, and central authorities in February/March of the following year.
5. It will cost the Ministry of Education of Honduras (a country I would place in the BoP category) about $2 million to run its National Assessment System per year, out of its national budget, supplemented by international monetary support.
6. It has cost about $40 million over 14 years, i.e. close to $3 million a year, to create and consolidate the infrastructure required for a National Assessment System to be productive and self-sufficient.
7. It has taken a decade to change educational culture in Honduras to understand, own, and use national content and performance standards and assessment instruments.
8. Gains have been sustained in both math and language pretty steadily since 2008.

In conclusion, there is a hierarchy of priorities that governs the usefulness of standardized assessment. It starts with content standards, upon which one develops a standards-aligned national assessment system used to provide in-grade feedback to students and teachers. A standards-aligned summative assessment system, together with performance scales, allows for performance data that is immediately used to estimate targets to guide instruction for the following academic year, and to provide subregionally sensitive input that is fed into central-level policy and strategy design. Participation in international assessment programmes, in my view, is either irrelevant or should rather be used to support locally established infrastructures and capacity.

Chapter 5
Educational Prosperity: an assessment strategy for supporting student learning in low-income countries

J. Douglas Willms

Introduction

Low- and middle-income countries have participated in large-scale international assessments since the 1970s. These have included mainstream studies, such as PIRLS, PISA, and TIMSS, as well as studies designed for low- and middle-income countries such as the Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación study (LLECE), the Literacy Assessment and Monitoring Programme (LAMP) and the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ). The contextual framework for these studies has embraced a multilevel 'school effects' paradigm, attempting to capture the most salient student, family, classroom, and school factors that explain student achievement.

In many respects, this paradigm has been successful. During this period, multilevel models were developed and refined, which provided a structure and approach for more accurate estimations of classroom and school effects (Aitkin and Longford, 1986; Goldstein, 1986; Raudenbush and Bryk, 1986; Raudenbush and Willms, 1995). Another advancement was the development of more sensitive measures of student outcomes and the factors associated with the context of the school, including parental involvement, teachers' expectations for achievement, teacher–student relations, and the disciplinary climate of the classroom (Brookover *et al.*, 1978; Gamoran, 1986, 1987; Ho and Willms, 1996; Pallas, 1988; Plewis, 1991; Slavin, 1990). Detailed studies of teachers' classroom practices found that the effective use of class time and structured and adaptive teaching were strongly associated with student learning (Scheerens, 1993; Slavin, 1994). Curriculum coverage and the content and pace of the curriculum also played a key role in student learning (Alexander, 1982; Barr and Dreeben, 1983; Dreeben and Gamoran, 1986; Lee and Bryk, 1989).

The international cross-sectional studies have provided evidence on the levels of students' reading and mathematics skills in low- and middle-income countries. However, in many reports, the analyses based on the 'school effects' paradigm have provided misleading results. The term 'school effects' refers to the effects that schools bring to students' learning through teachers' classroom instruction and other opportunities in the school setting (Raudenbush and Willms, 1995). The term 'learning' connotes a *change* in students' knowledge and skills over a period of schooling. The estimation of school effects in large-scale cross-sectional studies entails two assumptions: one is that students start school with similar levels of knowledge and skills, and the second is that the contemporary measures of various school resources and processes are an adequate proxy of the students' school experiences over the course of their school career. The causal claims that are often made from large-scale studies cannot be supported.

Aside from the issues associated with causation, the application of the school-effects paradigm in low- and middle-income countries is challenging for several reasons. The first and perhaps most important reason is that a large percentage of students in each country score at or near the floor of the achievement tests (Nonoyama-Tarumi and Willms, 2010). Consequently, there is little variation in student outcomes, either within or between schools. Second, the estimation of school effects also requires a sample of schools that vary in their school resources and processes, and a strategy for controlling for the SES of students and the mean SES of the school. This is problematic because measures of school resources and processes tend to be inter-correlated and strongly correlated with the average SES of the school. In cross-sectional studies such as PIRLS, TIMSS, and PISA, it is virtually impossible to isolate the school effects attributable to particular resources or processes. This problem is especially acute in low- and middle-income countries, not only because the correlations among the process and resource factors are stronger, but also because they are strongly correlated with the average SES of the school. Third, the most important driver of student learning – quality instruction – is not easily defined or measured. Moreover, quality instruction interacts with school resources in its effects on student outcomes; as Nonoyama-Tarumi, Hughes, and Willms (2015) noted, even the best teachers are unlikely to succeed without a certain level of material resources. As both factors are highly correlated with SES and with each other in low-income countries, one cannot estimate their effects (Nonoyama-Tarumi and Willms, 2010).

Willms (2016) set out a framework called Educational Prosperity, which calls for abandoning the school effects paradigm in favour of collecting rich data on a small set of developmental outcomes and the causal factors that drive these outcomes at several stages of children's development, from conception to adolescence. The approach is being used in a number of middle- and high-income countries, and has been embraced by the OECD and the countries participating in PISA for Development. Its use has implications for how large-scale assessments can be used to support student learning in low- and middle-income countries at the classroom and school levels, and how large-scale assessments can inform educational policy at the district or state levels.

The next section provides an overview of the Educational Prosperity model. The subsequent two sections discuss the monitoring requirements for two of the six phases of the Educational Prosperity framework. These sections include examples of assessment programmes that address some of the key challenges for conducting assessments that support learning in low- and middle-income countries. The paper concludes by setting out several criteria for an assessment strategy that can support student learning in low-income countries.

Educational Prosperity – a life-course approach

Educational Prosperity refers to the success of families, communities, and public institutions in developing children's cognitive skills and their social, emotional, physical, and spiritual well-being. The term 'prosperity' refers to the condition of experiencing success or thriving. The Educational Prosperity framework embraces a life-course approach, with key outcomes for six stages of development covering conception to adolescence (Willms, 2016). It includes a core set of outcomes, called 'Prosperity Outcomes', for each of the six stages of development, and a set of family, institutional, and community factors, called 'Foundations for Success', which drive these outcomes. The outcomes are considered universal in that they are key markers of child development.

Similarly, the Foundations for Success are universal in that a large body of research confirms that they are necessary conditions for success at each stage of development in low-, middle-, and high-income countries alike. Three criteria were considered in determining which factors to include as Foundations for Success: the factors must be potent, proximal, and pervasive. A 'potent' factor is one that has a strong correlation with an outcome or set of outcomes. For example, the quality of classroom instruction is arguably the most important driver of student outcomes

during the schooling period (Anderson, 2004; Rosenshine, 2010, Kyriakides, Christoforou, and Charalambous, 2013; Creemers and Kyriakides, 2006). A 'proximal' factor is close to the outcome – close in the sense that its relationship with the outcome is not mediated through some other factor. For example, the quality of classroom instruction has a direct, positive relationship on student outcomes, without any intervening factors. Principal leadership is also an important factor, and several studies have shown that it is correlated with student outcomes. However, it is not proximal because the 'effects' of principal leadership are mediated through the school-related foundations factors, namely inclusive context, quality instruction, learning time, material resources, and family and community support. Thus, a jurisdiction may allocate resources to improving principal leadership, but this would result in improved outcomes only if it led to improvements in one or more of the Foundations for Success. A 'pervasive' factor is positively correlated with a wide range of outcomes, although the strength of the correlation may vary with each outcome. For example, the effects associated with an 'inclusive school context' affect not only students' academic achievement, but also their educational attainment, their health and well-being, and their social, institutional, and intellectual engagement. Generally, the factors considered as Foundations for Success are also causal factors, based on evidence from several experimental studies.

Table 5.1 shows the Prosperity Outcomes and the Foundations for Success of the Educational Prosperity framework. It considers six periods of the life-course from conception to age 18: prenatal, early development (ages 0 to 2), pre-primary (ages 3 to 5), early primary (ages 6 to 9), late primary and lower secondary (ages 10 to 15), and upper secondary (ages 16 to 18). This classification does not necessarily mirror the structure of the school system in all countries, but it enables one to discuss in general terms the factors that affect children's development over this period of the life course.

The Educational Prosperity framework is based on a life-course approach, which considers children's development as the result of their personal characteristics, their actions, their culture, and the contexts in which they live (Mayer, 2009). The framework describes four processes by which success accumulates from one stage to the next (see *Figure 5.1*).

The outcomes at birth are affected by the Foundations for Success (striped arrow), which to some extent are biologically embedded (dotted arrow) through epigenetic processes. The age 2 outcomes are determined by a cumulative effect (white arrow), the Foundations for Success

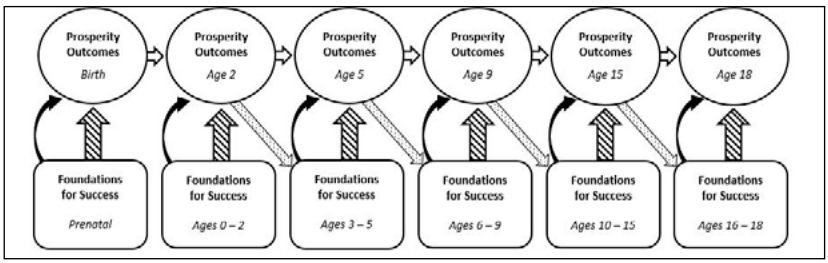
Table 5.1 The Educational Prosperity Framework

Prosperity Outcomes	Healthy pregnancy Healthy delivery	Language development Cognitive development Physical development	Awareness of self and environment Language development Cognitive development Physical development Social skills and approaches to learning	Reading literacy Numeracy Health and well-being Engagement	Academic achievement Educational attainment Health and well-being Engagement	Ethical citizen Leadership skills Health and well-being Communication and interaction skills
Family	Nutrition No exposure to toxins Mother's physical health Mother's emotional health	Breast-feeding and nutrition Mother's physical health Mother's emotional health Parenting skills Intra-family relations	Parenting skills Intra-family relations Family involvement	Parenting skills Intra-family relations Family involvement	Parenting skills Intra-family relations Family involvement	Parenting skills Intra-family relations Family involvement
Institution	**Health-care facility:** Prenatal care Primary health care	**Health-care facility:** Post-natal care Primary health care	**Pre-Schools:** Child-centered Goal-oriented Opportunities to socialize	**Schools:** Inclusive context Explicit teaching in: code-related skills and language skills Learning time Material resources	**Schools:** Inclusive context Quality instruction Learning time Material resources	**Schools:** Inclusive context Quality instruction Opportunities to learn career and life skills
Community	Social capital Resources	Social capital Resources	Social capital Resources	Social capital Resources	Social capital Resources	Social capital Resources
	Pre-Natal	**Early Development (Ages 0 – 2)**	**Pre-Primary (Ages 3 – 5)**	**Early Primary (Ages 6 – 9)**	**Late Primary and Lower Secondary (Ages 10 – 15)**	**Upper Secondary (Ages 16 – 18)**

Source: Willms, 2016.

associated with that stage (striped arrow), and an effect that is biologically embedded through the sculpting of the brain during critical periods (black arrow). The age 5 outcomes are also determined by cumulative effects, foundation effects, and biologically embedded effects. In addition, there is an institutional selection effect (dotted arrow), which leads to differential foundation effects for students with differing family backgrounds or ability. The outcomes at ages 9, 15, and 18 are affected by the same factors. The four processes are described below.

Figure 5.1 Four ways that success accumulates



Source: Willms, 2016.
Note: Biological embedding is represented by black arrows, Foundations for Success by striped arrows, cumulative effects by white arrows, and institutional selection effects by dotted arrows.

Biological embedding

Recent advances in neurobiology, molecular biology, and genomics have provided compelling evidence that children's early experiences interact with their genetic disposition in ways that affect brain development as well as other neurological and biological systems associated with healthy child development (Boyce, Sokolowski, and Robinson, 2012). Some of these biological processes are 'biologically embedded' during the prenatal period through epigenetic processes in which chemical signatures are attached to genes that predispose the child to either vulnerability or resilience (Boyce and Kobor, 2015).

Children are born with billions of neurons. During the course of early development, the neurons form connections called synapses in response to environmental stimuli. As this occurs, many of the neurons that are not being used are pruned away. This process of synapse formation and neuron pruning – the sculpting of the brain – is more rapid during certain critical periods of the first two or three years of life (McEwan and Schmeck, 1994; Cynader and Frost, 1999; Hertzman, 1999).

Foundations for Success

The second process involves the effects of the 'Foundations for Success'. As children develop, their outcomes are affected mainly by a small set of factors at each stage of development. For example, during the period from birth to age 2, children's development is affected by parents' engagement with the child and intra-family relations. From age 2 to age 5, their development is affected by these factors as well as by the quality of care at home and in early childhood centres. The selection of the foundation factors was based on theory and a large body of research that provides evidence of the effects of each factor on student outcomes. Hattie's (2009) monumental work was especially useful, as he synthesized the results from several meta-analyses of studies that assessed the effects of various school processes and resources on student outcomes.

Cumulative development

The third process is 'cumulative development'. Children develop their literacy skills in a cumulative way as they move from one stage to the next. The rate at which they develop these skills depends on the strength and duration of their exposure to the family, institution, and community factors that comprise the Foundations for Success. For example, a child's literacy skills at age 15 depend on his or her literacy skills at age 8, which is strongly affected by the quality of instruction the child received during the primary grades. The increase in the child's literacy skills from age 9 to 18 depend on the quality of instruction he or she received during the late primary and secondary school years.

Institutional selection

The fourth process is 'institutional selection'. When students are successful at one stage of development, their life-course may be altered if they are selected into certain classes, school programmes, or schools. For example, children who have strong reading and language skills are more likely to be streamed into classes or school programmes where they benefit from positive peer interactions, a higher quality of instruction, and other factors that enable them to develop their skills at a faster pace. Children who experience learning difficulties at a particular stage are more likely to be streamed into lower-ability classes and have less access to the factors that improve their skills.

An assessment strategy to support early learning: the Early Years Evaluation in Uruguay

The Early Years Evaluation–Teacher Assessment (EYE-TA; see *Box 5.1*) has been used to assess the pre-literacy skills of all preschool children aged 4 and 5 in Uruguay (Willms, 2011). Its implementation entailed a pilot study in one of Uruguay's 19 departments in 2014, extension to two departments in 2015, extension to 10 departments in 2016, and a country-wide implementation in 2017. The EYE-TA is administered for all 4- and 5-year-old children at the beginning of the school year and with a subset of vulnerable children near the end of the school year. Results are reported at multiple levels: individual child, class, school, department, and country. Over the course of the school year, teachers receive information on classroom activities designed to strengthen children's development in each of the five domains. At higher levels of aggregation, the data are used to assess the extent of inequalities among various subpopulations and to inform decisions about the allocation of resources and educational policy. The Ministry assured teachers that the data would not be used for holding schools or teachers accountable. Instead, the model is one based on diffusion of best practice.

State-mandated assessments can be grown to scale very quickly, but these assessments are not necessarily embraced by teachers, nor do they automatically lead to improvement in student outcomes. In many countries, teachers and their unions resist the introduction of new assessments. Growing an assessment to scale from the bottom up has its own challenges, and tends to be a slower process. Lopez-Garcia (2016) noted four aspects of the implementation of the Early Years Evaluation in Uruguay that contributed to its success: (1) a rigorous examination of the psychometric properties of the assessment; (2) modification of the assessment to fit the cultural context, after consultation with curriculum experts and teachers; (3) a gradual implementation that allowed for identifying and addressing emerging issues; and (4) the involvement of teachers in the development of the implementation process and its use for best practice. As this process was unfolding, the government conducted reviews, each time asking whether the content of the assessment was relevant to kindergarten children's development and whether it contributed to improvements in their teaching practices.

After contextualization, the EYE-TA had strong psychometric properties, with a factor structure showing clear loadings for every item on their intended domains, and reliability coefficients ranging from 0.90

to 0.93. Modification of the instrument involved changes at the item level, generally associated with the fit with the curriculum or differing meanings of particular Spanish words. The involvement of teachers in this process was invaluable. The implementation process took three years before the national roll-out, which, compared with most interventions, is very fast. The Learning Bar's current contextualization and implementation process includes a two-year development period, with national or provincial implementation in the third year.

Box 5.1. The Early Years Evaluation (EYE)

The EYE is an assessment framework used to identify the developmental skills of children aged 3–6 years as they prepare for and make the transition to formal schooling (www.earlyyearsevaluation.com/index.php/en/). The tools assess skills in five developmental domains: Awareness of Self and Environment, Cognitive Skills, Language and Communication Skills, Physical Development, and Social Skills and Approaches to Learning. These skills are research-based predictors of children's later reading achievement and are consistent with the framework for school readiness currently used by UNICEF and the National Education Goals Panel. The assessment is being used widely across Canada as well as in a number of Latin American countries.

The EYE consists of two complementary tools that help educators monitor the overall development of children as they prepare for and transition to school. The EYE-Direct Assessment (EYE-DA) is a play-based, interactive assessment conducted by a trained evaluator. It requires about 45 minutes to complete. The data are entered into an online scoring application, and a report for the child is immediately generated. The EYE-Teacher Assessment (EYE-TA) provides a systematic framework that teachers can use to structure their observations and informal assessments. It is typically used by kindergarten teachers to provide them with formative, instructionally relevant information. At the beginning of the school year, after a period of two to three weeks of observation, teachers report on each child's skill development using an online data collection system. As with the EYE-DA, the results at the individual child and classroom level are immediately generated. The child and classroom reports are used in combination with other teacher assessments to identify students who are encountering difficulty. The classroom reports include a summary measure, Responsive Tiered Instruction (RTI), which indicates the *type and amount* of support required for each child.

Source: Willms, 2011.

The test of whether an assessment strategy supports learning in low-income countries is whether it leads to improvements in student performance from one period to the next. This requires changes in teachers' day-to-day classroom practices. The EYE assessment in Uruguay has provided strong evidence of the preschool programme's positive impact, more than would be expected than if the children were not attending a pre-school. Uruguay is classified as an upper middle-income country with a gross national income (GNI) of $15,720

in 2015 (World Bank, 2017). However, the contribution of regional departments to its GNI varies considerably, and the EYE implementation has been successful across all departments, including those with a relatively low contribution. The external evaluation conducted by the Uruguayan Ministry found that 89 per cent of teachers felt that an evaluation such as the EYE was necessary as it helped them plan their activities, and led to positive changes in their day-to-day practice (ANEP, 2017). Five factors have contributed to its successful implementation:

1. It has strong psychometric properties.
2. It is population-based, involving all preschool and kindergarten teachers conducting an assessment of their students.
3. The assessment is contextualized to meet the needs of Uruguayan teachers.
4. It is linked to an intervention that enables teachers to improve their practice.
5. The reports and use of the data position the EYE as a leading indicator, providing teachers with instructionally relevant information, rather than a trailing indicator used to hold teachers accountable.

An assessment strategy to support literacy skill development: the Confident Learners literacy programme for indigenous students

The challenge for all primary school teachers, in low-, middle- and high-income countries alike, is providing instruction at the appropriate level for students with a wide range of pre-literacy skills. A key finding of EYE analyses of children's skill development in Latin America is that when children are nearing the end of their kindergarten year, set to begin Grade 1, their pre-literacy skills span at least four years. About 70 per cent of children have skills that range from three-quarters of a year behind to a full year ahead of the norms for their age. However, about 22 per cent have skills that are between one-and-a-half years to three-quarters of a year behind their same-age peers, and a further 8 per cent have skills that are more than one-and-a-half years lower than their same-age peers. We expect that the levels of skills would be lower and the range would be greater in low- and middle-income countries (van der Berg *et al.*, 2016). This has profound implications for children's progression on the pathway, especially regarding the likelihood of their making a successful transition to school and eventually learning to read fluently and with confidence by the end of Grade 3.

The problem is exacerbated by grade-based formal curricula, which call for an inordinate number of skills to be taught in Grade 1. The literacy programme Confident Learners was designed to meet the needs of indigenous students in Canada, which include First Nations, Inuit, and Métis students. During the course of development of the programme, three literacy experts amassed the curricula from all Canadian provinces, some US states, and from Australia and the United Kingdom (Willms, 2016; see *Box 5.2*). They set out a model describing the skills at a micro-level that are required for decoding words and developing the language skills needed for learning to read during the primary school years. The model resulted in over 300 'coding' skills and 300 'language' skills covering the period from kindergarten to Grade 3.

An unexpected finding of this research was that the distribution of the coding skills spiked at the beginning of Grade 1 and remained high through to the middle of Grade 2; thereafter, it fell precipitously. Fifty-seven per cent of skills covered in the four-year period were to be taught from the beginning of Grade 1 through to the first half of Grade 2. We refer to this phenomenon as the 'reading mountain'.

Children who start school with skills that are one or more years below the norm have very little chance of traversing the reading mountain and emerging as proficient readers by the end of Grade 2 or 3. In the 2011 PIRLS study, for example, only 38 per cent of students in Columbia and 28 per cent in Indonesia reached the PIRLS intermediate benchmark when tested in Grade 4. The results are worse in the rural areas of these countries and in low-income countries.

A key element of Confident Learners is an approach to instruction and assessment that emphasizes children's progress on the 'pathway to literacy success'. The act of reading is complex and relies upon a combination of pre-literacy skills. Reading research has been successful in identifying a number of these crucial precursors to literacy development, and has provided considerable evidence of effective methods to utilize during early reading instruction (USA, 2008). The widely recognized 'Simple View of Reading' maintains that successful reading acquisition depends upon two complementary components: code-related skills and language skills (Gough and Tunmer, 1986). For a child to become a successful reader, he or she must not only master the ability to accurately decode written words, but also understand the meanings of words and how they combine in phrases, sentences, and paragraphs.

Box 5.2. Confident Learners

Confident Learners is a whole-school literacy programme aimed at improving the literacy skills of Indigenous children during the primary grades (http://confidentlearners.com/index.php/en/). It has five core elements:

- a "pathway approach" to instruction and assessment based on the science of literacy which is linked explicitly to teaching activities that support Indigenous language and culture, and aligned with efforts to increase children's language skills in their native language;
- a professional development programme for teachers aimed at increasing their knowledge of the science of literacy skill development, assessment practices, and high-yield teaching strategies;
- a formative and summative assessment programme which enables teachers to monitor each child's progress as they acquire literacy skills during the primary grades;
- a bank of over 500 fun, engaging, and culturally-relevant learning activities, with an online application that allows teachers to plan their lessons to meet students' individual needs; and
- a family and community literacy programme aimed at strengthening families' contributions to their children's literacy development.

Its implementation and ongoing development is supported by an Indigenous Advisory Circle and a team of 12 researchers, programme developers and educators at The Learning Bar Inc.

The development of Confident Learners entailed amassing curricula from the 10 Canadian provinces, as well as curricula from the UK, Australia, and the US. Three literacy experts worked with First Nations Educators to develop an instructional framework that includes 20 modules or 'steps' for coding skills and 20 steps for language skills. Each step has 14 skill-based objectives. The programme also includes an assessment framework with short one-on-one assessments to assess each child's mastery of the skills comprising each step. The professional development programme for teachers includes 16 modules relevant to four elements of student learning: quality instruction, student engagement, classroom and school context, and learning time. Each module is based on recent research relevant to these four elements. For example, a module called 'high-yield teaching strategies' focuses on structured teaching and ways teachers can incorporate the approach in their classroom. The training programme for parents and community members has three aims: increasing student attendance, ensuring each child is read to at least 15 minutes every day, and enabling parents to engage with their children in literacy activities at home.

Confident Learners is implemented over 30 weeks during the school year. The implementation period is comprised of six 'sprints' of five weeks each. During the first sprint, teachers assess each child using a placement test that determines where each child is on the literacy pathway. For example, a child in kindergarten may be at step two in coding and step three in language. However, some children begin school with a well-developed set of early literacy skills, and may be on step 10 or higher in either coding or language skills. Children who are in Grades 2 and 3 tend to have a wide range of skills. Teachers enter children's placement test results into the Confident Learners application, which provides a map of where each child is on the literacy pathway. The application allows teachers to group their children into small groups based on their level of skill development, and then links the groups to the appropriate learning activities

for each group. Teachers then build their lesson plans for the upcoming week or for an entire five-week sprint. During the last week of the sprint, teachers conduct assessments to discern whether students have mastered the objectives in one or more of the steps. The results are entered into the application, allowing teachers to track each child's progress.

Source: Willms, 2016.

Confident Learners requires teachers to make a shift from teaching a grade-based curriculum to teaching the specific skills that children need to improve their literacy skills. A teacher is no longer a 'Grade 1 teacher' or a 'Grade 2 teacher' charged with teaching the curriculum for a particular grade. Instead, he or she has a singular focus on enabling every child to move forward another step on the literacy pathway. Teachers using the programme also have to shift their practice in another way: they need to adopt a 'skills-based' approach instead of a 'deficit approach'. Rather than focusing on what children cannot do and trying to identify deficits, such as learning or intellectual disabilities, teachers take stock of what children can do and focus on the learning activities that enable them to take the next step on the pathway. Teachers are able to make these shifts as their work is supported with ongoing professional development, as well as support from the principal and a literacy lead in the school. The efforts of school staff are also supported by parents and the wider community.

The Learning Bar's implementation of Confident Learners is in its early stages. However, the research and feedback from teachers and principals provide three lessons relevant to the usefulness of assessment in low- and middle-income countries:

1. The assessments need to be skill-based, not grade-based, and cover a wide range of skills.
2. The assessments need to be linked to a professional development programme for teachers, which strengthens their knowledge of literacy assessment and high-yield teaching strategies.
3. Children are more likely to improve their skills if the assessments are linked to engaging and culturally relevant learning activities.

Implications for assessment in low- and middle-income countries

The application of the Educational Prosperity model in low- and middle-income countries calls for a relentless focus on building the 'Foundations for Success'. The emphasis in this paper has been on how assessment can help improve quality instruction, which needs to begin during the preschool years. The most important transition for students in all

countries is the transition from 'learning-to-read' to 'reading to learn'; that is, children need to learn to read with confidence during the primary grades, because after Grade 3, the emphasis shifts towards understanding curricular materials in subject domains such as health, social studies, and science. Students are expected to learn the languages of subject domains and use that language to think critically, solve problems, and create new knowledge. Moreover, the demands for strong reading skills increase as students make their way into the higher grades. Students who lack the fundamental reading skills fall further and further behind.

Many governments of low- and middle-income countries have invested heavily in developing national assessments of student learning at particular grades. In addition, many countries are participating in international assessments. These assessments can serve to establish standards, assess the extent of inequalities among various sub-populations, inform educational policy, and provide a framework for basic or theoretical research. However, they have little impact on changing classroom practice.

The assessments discussed in this paper – the Early Years Evaluation and Confident Learners – situate assessment as a leading indicator, aimed at identifying students who need extra support, guiding classroom practice, increasing student and teacher engagement, and involving parents in meaningful ways. They have proven to be effective in low- and middle-income jurisdictions. If assessments are to support student learning in low- and middle-income countries, they need to be census-based, contextualized with the involvement of teachers, skill-based rather than curriculum-based, and linked explicitly with teacher professional development and a classroom intervention.

References

Aitkin, M.; Longford, N. 1986. 'Statistical modeling issues in school effectiveness studies'. In: *Journal of the Royal Statistical Society, Series A*(1), 1–43.

Alexander, K.L. 1982. 'Curricula and coursework: a surprise ending to a familiar story'. In: *American Sociological Review*, *47*(October), 626–640.

Anderson, L.W. 2004. *Increasing teacher effectiveness,* 2nd edn. Paris: IIEP-UNESCO.

Barr, R.D.; Dreeben, R. 1983. *How schools work.* Chicago, Ill.: University of Chicago Press.

Boyce, W.T.; Kobor, M.S. 2015. 'Development and the epigenome: the "synapse" of gene–environment interplay'. In: *Developmental Science, 18*(1), 1–23.

Boyce, W.T.; Sokolowski, M.B.; Robinson, G.E. 2012. 'Toward a new biology of social adversity'. In: *Proceedings of the National Academy of Sciences, 109*(2), 17143–48.

Brookover, W.B.; Schweitzer, J.H.; Schneider, J.M.; Beady, C.H.; Flood, P.K.; Wisenbaker, J.M. 1978. 'Elementary school social climate and school achievement'. In: *American Educational Research Journal*, 15, 301–318.

Creemers, B.P.M.; Kyriakides, L. 2006. 'Critical analysis of the current approaches to modelling educational effectiveness: the importance of establishing a dynamic model'. In: *School Effectiveness and School Improvement, 17,* 347–366.

Cynader, M.S.; Frost, B.J. 1999. 'Mechanisms of brain development: neuronal sculpting by the physical and social environment'. In: D. Keating and C. Hertzman (Eds), *Developmental health and the wealth of nations: social, biological and educational dynamics* (pp. 153–184). New York: Guilford Press.

Dreeben, R.; Gamoran, A. 1986. 'Race, instruction, and learning'. In: *American Sociological Review*, 51, 660–69.

Gamoran, A. 1986. 'Instructional and institutional effects of ability grouping'. In: *Sociology of Education, 59*(October), 185–198.

——. 1987. 'The stratification of high school learning opportunities'. In: *Sociology of Education, 60*(July), 135–155.

Goldstein, H. 1986. 'Multilevel mixed linear model analysis using iterative generalized least squares'. In: *Biometrika, 71*(1) 43–56.

Gough, P.B.; Tunmer, W.E. 1986. 'Decoding, reading and reading disability'. In: *Remedial and Special Education,* 7(1), 6–10.

Hattie, J. 2009. *Visible learning: a synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.

Hertzman, C. 1999. 'The biological embedding of early experiences and its effects on health in adulthood'. In: *Annals of the New York Academy of Sciences,* 896, 85–85.

Ho, E.; Willms, J.D. 1996. 'The effects of parental involvement on eighth grade achievement'. In: *Sociology of Education, 69*, 126–141.

Kyriakides, L.; Christoforou, C.; Charalambous, C.Y. 2013. 'What matters for student learning outcomes: a meta-analysis of studies exploring factors of effective teaching'. In: *Teaching and Teacher Education, 36*, 143–152.

Lee, V.E.; Bryk, A.S. 1989. 'A multilevel model of the social distribution of high school achievement'. In: *Sociology of Education, 62*, 172–192.

Lopez-Garcia, A.Y. 2016. 'Steps towards a national implementation of the Early Years Evaluation in Uruguay: actors and actions'. Paper for Inaugural Symposium of the Comparative and International Education Society, November.

Mayer, K.U. 2009. 'New directions in life course research'. In: *Annual Review of Sociology, 35*, 413–33.

McEwan, B.S.; Schmeck, H.M. 1994. *The hostage brain.* New York: Rockefeller University Press.

Nonoyama-Tarumi, Y.; Hughes, K.; Willms, J.D. 2015. 'The role of family background and school resources on elementary school students' mathematics achievement'. In: *Prospects, 45*(3), 305–24.

Nonoyama-Tarumi, Y.; Willms, J.D. 2010. 'The relative and absolute risks of disadvantaged family background and low levels of school resources on student literacy'. In: *Economics of Education Review, 29*(2), 214–224.

Pallas, A. 1988. 'School climate in American high schools'. In: *Teachers College Record*, 89(4), 541–553.

Plewis, I. 1991. 'Using multilevel models to link educational progress with curriculum coverage'. In: S.W. Raudenbush and J. D. Willms (Eds), *Schools, classrooms, and pupils: international studies of schooling from a multilevel perspective* (pp. 149–166). San Diego, Calif.: Academic Press.

Raudenbush, S.W.; Bryk, A.S. 1986. 'A hierarchal model for studying school effects'. In: *Sociology of Education, 59*, 1–17.

Raudenbush, S.W.; Willms, J.D. 1995. 'The estimation of school effects'. In: *Journal of Educational and Behavioral Statistics, 20*(4), 307–335.

Rosenshine, B. 2010. 'Principles of instruction: International Academy of Education'. Geneva: UNESCO International Bureau of Education.

Scheerens, J.; Vermeulen, C.; Pelgrum, W.J. 1989. 'Generalizability of instructional and school effectiveness indicators across nations'. In: *International Journal of Educational Research, 13*, 789–799.

Scheerens, J. 1993. 'Effective schooling: research, theory and practice'. In: *School Effectiveness and School Improvement, 4*(3), 230–235.

Slavin, R.E. 1990. 'Achievement effects of ability grouping in secondary schools: a best evidence synthesis'. In: *Review of Educational Research, 60*, 471–499.

——. 1994. 'Quality, appropriateness, incentive, and time: a model of instructional effectiveness'. In: *International Journal of Educational Research, 21*(2), 141–158.

Uruguay. 2017. Administración Nacional de Educación Pública (ANEP). *Encuesta a maestros sobre la Evaluación Infantil Temprana en Uruguay.* Montevideo: Ministry of Education.

USA. 2008. National Early Literacy Panel. *Developing early literacy: report of the National Early Literacy Panel.* Washington DC: National Institute for Literacy. Retrieved from: http://lincs.ed.gov/publications/pdf/NELPReport09.pdf.

Van der Berg, S.; Spaull, N.; Wills, G.; Gustafsson, M.; Kotzé, J. 2016. *Identifying binding constraints in education: Synthesis report for the Programme to support Pro-Poor Policy Development (PsPPd).* Stellenbosch, South Africa: Department of Economics, University of Stellenbosch.

Willms, J.D. 2011. 'Early years evaluation'. Retrieved from: www.earlyyearsevaluation.com

——. 2016*a*. 'Confident learners.' Retrieved from: http://confidentlearners.com/index.php/en/

——. 2016*b*. *Educational prosperity*. Fredericton, N.B.: The Learning Bar.

World Bank. 2017. World Bank country and lending groups. Accessed 13 June 2017: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Commentary

Sarah Howie

The chapter by Willms presents a thought-provoking theoretical framework – 'the Educational Prosperity model' – to support student learning in low-income countries, which is both holistic and comprehensive. Most of the frameworks developed internationally to date have been based upon research in developed countries (Howie, Scherman, and van Staden, 2016) and are insufficient for explaining the variance across and within schools in low- to middle-income countries. While the chapter provides an interesting theoretical model, it does not yet deal with the realities facing low- (and middle-) income countries in terms of its feasibility and application in such settings.

The chapter highlights the importance of international assessments in providing insights and important information for policy-makers to take action, while critiquing some limitations. Comprehensive analyses have revealed considerable differences between high- and low-income countries (e.g. Howie and Plomp, 2006; Martin and Mullis, 2013) and associated factors (Nilsen and Gustaffson, 2016). Not noted in the paper are the developments in PIRLS[15] (prePIRLS and PIRLS Literacy) which address the bottom of the pyramid and measurement problems of previous studies (Mullis *et al.*, 2009; Howie, 2015) in low- and middle-income countries. The paper highlights the fact that in wealthier countries most of the variance in performance is found within schools, yet there is evidence to suggest that this is not the case in countries that have significant inequality in education provisioning and quality, where most of the variance is found between schools (Howie, 2003).

The assessment strategy example from Uruguay appears similar to others, such as International Performance Indicators in Primary Schools currently being trialled in other middle-income countries (Howie *et al.*, 2016) and reveals a number of challenges that would face others in a low-income context. The Literacy Programme proposed appears noteworthy, but the evidence for its viability is largely based upon Canada, the UK, and the USA. It is not clear how it could be adapted to more disadvantaged environments in low-income countries where parents may not be able to read to children; where teachers may

15. Progress in International Reading Literacy Study 2011 and 2016 included new assessments for lower-achieving students (see Mullis *et al.*, 2012, 2017).

be poorly trained; or where classrooms are very large, with more than 50 learners. Nor is it clear how these types of factors would affect such a programme. These types of factors are currently being addressed by literacy programmes such as Read, Molteno, and PRAESA (amongst others) in South Africa.

Finally, the criteria for assessments that can support learning and the components of the framework are important contributions to the ongoing quality of education debates (Howie, 2012). One possible amendment to the proposed framework would be to include the quality of instruction related to pre-literacy and pre-numeracy skills from 3–5 years of age onwards. Given the considerable variance between schools in many countries, one possible intervention is to ensure that all (especially disadvantaged) children are afforded the opportunity to get a solid start at a pre-primary level and that the skills and knowledge are facilitated at that earlier stage. Pre-primary also delivers instruction, albeit often integrated within play. The challenge is that low-income countries today are struggling to deliver on their compulsory primary and secondary school education agenda and therefore, pre-primary is not yet a priority in many settings, despite the obvious need. The same is true for the assessment strategy, where those who may need it most (those at the BoP), can least afford it (Howie, 2016), and therefore the implementation challenges have to be addressed before such a model is viable.

References

Howie, S.J. 2003. 'Language and other background factors affecting secondary pupils' performance in mathematics in South Africa'. In: *African Journal for Mathematics, Science and Technology Education*, *7*, 1–20.

——. 2012. 'High-stakes testing in South Africa: friend or foe? Special issue: high stakes testing – value, fairness and consequences'. In: *Assessment in Education, 19*(1), 81–98.

——. 2015. 'What do the IEA studies mean for developing countries education systems and educational research?' Keynote address for 6th IEA International Research Conference, Cape Town, South Africa, 24–26 June.

——. 2016. 'Assessment for political accountability or towards educational quality?' Keynote addres for International Association for Educational Assessment (IAEA) Conference, Cape Town, South Africa, 21–26 August.

Howie, S.; Combrinck, C.; Tymms, P.; Merrell, C. 2016. 'What children know and can do when they start school in the Western Cape'. Durham, UK: Centre for Evaluation and Monitoring, University of Durham.

Howie, S.J.; Plomp, T. (Eds). 2006. *Contexts of learning mathematics and science: lessons learned from TIMSS*. London: Routledge Falmer.

Howie, S.J.; Scherman, V.; van Staden, S. 2016. 'Frameworks for the analysis of monitoring data related to achievement'. In: V. Scherman, R. Bosker, and S.J. Howie (Eds), *Monitoring the quality of education in schools*. Rotterdam, Netherlands: Sense.

Martin, M.O.; Mullis, I.V.S. 2013. 'TIMSS and PIRLS 2011: relationships amongst reading, mathematics and science achievement at the fourth grade – implications for early learning'. Chestnut Hill, Mass.: Lynch School of Education, Boston College. https://timssandpirls.bc.edu/timsspirls2011/downloads/TP11_Relationship_Report.pdf

Mullis, I.V.S.; Martin, M.O.; Kennedy, A.M.; Trong, K.L.; Sainsbury, M. 2009. *PIRLS 2011 Assessment Framework*. Chestnut Hill, Mass.: Lynch School of Education, Boston College. https://timssandpirls.bc.edu/pirls2011/downloads/pirls2011_framework.pdf

Mullis, I.V.S.; Martin, M.O.; Foy, P.; Drucker, K.T. 2012. *PIRLS 2011 international results in reading*. Chestnut Hill, Mass.: Lynch School of Education, Boston College. https://timssandpirls.bc.edu/pirls2011/downloads/P11_IR_FullBook.pdf

Mullis, I.V.S.; Martin, M.O.; Foy, P.; Hooper, M. 2017. *PIRLS 2016 international results in reading*. Chestnut Hill, Mass.: Lynch School of Education, Boston College. http://timssandpirls.bc.edu/pirls2016/international-results/wp-content/uploads/structure/CompletePDF/P16-PIRLS-International-Results-in-Reading.pdf

Nilsen, T.; Gustaffson, J-E. (Eds). 2016. *Teacher quality, instructional quality and student outcomes: relationships across countries, cohorts and time*. IEA Research for Education 2. New York: Springer. DOI 10.1007/978-3- 319-41252-8_2.

Commentary

Marlaine Lockheed

This chapter critiques the 'school effects' literature, proposes an alternative life-course 'Educational Prosperity' model, provides examples of the model as implemented, and concludes with implications for assessments in low- and middle-income countries. The model has many commendable features, including a 'relentless focus' on leading indicators of literacy and numeracy. The chapter serves as a valuable reminder of a main purpose of schooling, but it does not address the implementation challenges or resource needs associated with the model.

To justify the new model, Willms summarizes many familiar shortcomings of cross-sectional studies, both generally and in low- and middle-income countries. The critique does not apply to all assessments. Some international large-scale assessments, such as PISA and TIMSS, may utilize tests that are insensitive to student performance at the low end of the scale and survey instruments that incompletely assess school and classroom resources, processes, and instructional quality (Lockheed *et al.,* 2015). But regional assessments, such as the LLECE, SACMEQ, and PASEC,[16] utilize student outcome measures and contextual questionnaires designed to reflect regional standards (Lockheed, 2012, 2016). And the World Bank's Service Delivery Indicator (SDI) studies – implemented in six sub-Saharan African countries – include measures of basic literacy, teacher knowledge, and teaching practice (Amin, Das, and Goldstein, 2008).

The chapter summarizes a life-course approach to learning and the requisites for success, and discusses how this approach can inform assessments and interventions in low-income countries. It provides two examples – the Early Years Evaluation in Uruguay (an upper middle-income country where 97 per cent of 5-year-olds are in kindergarten) and the 'Confident Learners' literacy programme for indigenous children in three high-income countries. Both involve proprietary software[17] and complex implementation strategies. I question whether these approaches are affordable and implementable in low-income

16. Programme d'analyse des systèmes éducatifs de CONFEMEN, Southern and Eastern Africa Consortium for Monitoring Education Quality.

17. Available from The Learning Bar www.thelearningbar.com/?lang=en.

countries where fewer than 20 per cent of children attend preschool and where overall policy implementation is weak.

The description of the 'prosperity outcomes' model in Uruguay reveals some of the challenges. It required: (a) a trained evaluator who individually administered a 45-minute 'play based interactive assessment' to all 4- and 5-year old children, (b) data entry via an 'online scoring application', (c) an individual report immediately generated for each child, as well as (d) continuing follow-up. Another tool is available to assist teachers in individually observing children. In low-income settings – where less than one-fifth of 4–5 year-olds-attend pre-school, first grade class sizes of over 100 students are common, internet coverage is low, and many schools lack electricity – using either of these tools presents a sizeable implementation challenge.

Ministers of education in low-income countries are aware that children are not learning; early grade reading assessments such as EGRA, UWEZO, and ASER[18] provide this information. Ministers also understand many of the requisites for learning. But their main challenges are ensuring that resources are aligned with desired outcomes and implementing policies related to the financing, deployment, and monitoring of the resources. To utilize these assessment tools in the poorest countries, where the need is greatest, may exceed both the budgetary and implementation capacity of these countries.

References

Amin, S.; Das, J.; Goldstein, M. (Eds). 2008. *Are you being served? New tools for measuring service delivery.* Washington DC: World Bank.

Lockheed, M. 2012. 'Policies, performance and panaceas: the role of international large-scale assessments in developing countries'. In: *Compare: a Journal of Comparative and International Education, 42*(3).

——. 2016. 'Measures that matter: learning outcome targets for Sustainable Development Goal 4: an examination of national, regional and international learning assessments', paper commissioned for *Education for people and planet: creating sustainable futures for all,* Global Education Monitoring Report 2016. Paris: UNESCO. http://unesdoc.unesco.org/images/0024/002458/245842E.pdf

18. Early Grade Reading Assessment, Annual Status of Education Report (ASER), UWEZO ('capability' in Kiswahili, used in Kenya, Tanzania and Uganda).

Lockheed, M.; Prokic-Bruer, T.; Shadrova, A. 2015. *The experience of middle-income countries participating in PISA 2000–2015.* Paris: PISA/OECD Publishing.

Chapter 6
Localized assessments can enhance our understanding of learning at the bottom of the pyramid: insights from India

Shaher Banu Vagh and *Dropti Sharma*

Assessments have the potential to not only provide information about the status of learning, but also help guide teaching–learning processes and set goals. In fact, well-designed assessments can help underscore the importance of domain-specific sub-skills. To be well designed, assessments of literacy acquisition need to be sensitive to local and contextual needs by accounting for the orthography-specific characteristics of the language of instruction, programme or curriculum requirements, and the backgrounds of children for whom they are designed. But all of this can be effectively galvanized only after considering common operational constraints, such as contextual capabilities, testing time demands, and ability to generate timely feedback for formative or summative purposes. An additional important consideration in all of this is the 'purpose' of assessment, essentially the intended inference and use of test scores. This is particularly relevant for children who are from marginalized populations, or at the 'bottom of the pyramid', as these children are more likely to perform poorly on standardized assessments that are usually not adapted to inform their learning needs.

This paper discusses a collaborative, action research project conducted by Pratham in conjunction with the programme's teachers, which sought to develop and evaluate a literacy and numeracy programme for children from low socio-economic backgrounds in Grades 1–3 attending government schools in Allahabad, India. It highlights some of the strengths and challenges of localized assessments as implemented in this collaborative, longitudinal programme and discusses the role of assessments in enhancing the teaching–learning process.

Conceptual framework

The conceptual underpinnings of the programme, drawing from an emergent literacy paradigm (Teale and Sulzby, 1986) and research on the *akshara* (Indian) languages (see Vagh, Nag, and Banerji, 2017),

are that reading and writing skills develop through active engagements with print, during activities mediated and scaffolded by adults, as well as through direct and explicit instruction, that takes into account the contextual realities of the children's home and community environments. For instance, creating a print-rich classroom environment was seen as a key goal given that participating children's home environments provided limited print exposure.

The akshara languages: an overview of their orthography, phonology, and literacy acquisition research

In the orthographies representing Indian languages, the basic orthographic unit – the akshara –represents a syllable, with constituent parts of the akshara encoding phonemic information. The akshara writing system is most commonly termed an *alphasyllabary* since the akshara can represent a phoneme or a syllable (Salomon, 2000). Akshara can be simple, such as the primary form of vowels, or consonants with an inherent vowel, or complex, like consonant-vowel ligature pairs, conjoint consonants, and consonant clusters. The consonants with the full set of accompanying vowel ligatures are presented in a matrix, which in Hindi is called the *barakhadi*. Vowels have two representations, a primary form, which mostly occurs in syllable initial positions, and a secondary form that appears as a diacritic that can be ligatured to the top, bottom, left, or right of the consonant base, resulting in a non-linear orthographic representation. Additionally, consonants within clusters are also modified by halving, and placed either linearly or stacked vertically. Given the versatility of the akshara, singleton akshara may amount to several hundred, resulting in an 'extensive' orthography (Nag, 2007, 2013). An additional distinct feature of the orthography is that the phonemic units or markers can be deconstructed for the majority of the akshara, with the exception of the inherent vowel and a few conjoint consonants (such as the Hindi /ksh/).

Research on the akshara languages suggests that while simple akshara are acquired rather quickly in Grades 1 and 2, the acquisition of more complex akshara is slower, and can well extend into Grade 4 (Nag, 2007). Factors contributing to the slower emergence of the complex akshara are attributable to their visual (non-linear composition) and phonological (multiple phonemic markers) complexity. However, the lower frequency of occurrence of many complex akshara coupled with limited availability of print resources may further impede the learning of complex akshara and extend the 'learning to read' phase. A noteworthy finding is that learning

complex akshara corresponds with increased sensitivity to phonemes, suggesting that analytical skills that allow children to deconstruct an akshara into its constituent phonemic units aid or may be reciprocally associated with reading gains. In keeping with the alphasyllabic nature of the akshara orthography, both phonemic and syllabic awareness skills seem to play a critical role in the 'learning to read' process, well into the middle school years (Nag and Snowling, 2012).

Research on spelling development highlights that (a) complex akshara are as difficult to spell as they are to read relative to the simple akshara, and (b) the presence of mismatched phonology-orthography mappings has greater consequences for writing than for reading. An important consideration is the widely prevalent phenomenon of diglossia, i.e. the prevalence of a vernacular community dialect or language that differs from the school or standard variety of the same language. For instance, regional speakers in Bihar tend to not differentiate between the phonemes /s/ and /sh/, which has implications for spelling as well as for reading.

Comprehension studies indicate that despite an extended 'learning to read' process, young readers are able to comprehend text well before the full acquisition of complex akshara, which suggests the strong scaffolding role of lexical knowledge (Kannada: Nag and Snowling, 2012; Gujarati: Patel, 2004). In addition, although fluency is a significant correlate of comprehension in Hindi, the unique contribution of context fluency (4 per cent) relative to list fluency (2 per cent) is small when compared with English (41 per cent vs. 1 per cent), and accuracy measures remain significant predictors even after accounting for oral reading fluency, suggesting that fluency is a necessary but not sufficient condition for reading with meaning (English: Jenkins *et al.*, 2003; Hindi: Vagh and Biancarosa, 2012). In sum, the research suggests that good comprehenders demonstrate better reading accuracy, phonological skills, and a higher lexical repertoire.

The ensuing implications and challenges for reading instruction and learning are:

- The acquisition of simple akshara and the consonant-vowel ligature pairs (the barakhadi) are important in the early grades. Key challenges for young readers centre on discriminating between phonetically similar symbols (e.g. /s/ and /sh/) and phonemic ligature markers for some short and long vowel pairs that differ only in their orientation (e.g. /i/ and /ii/).

- Explicit instruction that fosters syllabic- and phonemic-level analytical skills is likely to facilitate the acquisition of complex akshara, which in turn accelerates the pace of the 'learning to read' process.
- Oral language proficiency is critical for successful decoding and text comprehension.
- Access to a print-rich environment has many benefits. It facilitates stable phonology–orthography linkages by providing exposure to diverse akshara, and it helps readers gain an understanding of the context-dependent rules of akshara formation. Print engagements also contribute favourably to vocabulary knowledge, knowledge of syntax, and world knowledge, all of which in turn enhance reading comprehension.
- Instructional practices need to account for and build upon home and community dialects/languages.

Given the conceptual framework and the research base, the broad set of literacy goals drafted for the programme were that children should be encouraged to:

- communicate, share ideas, and orally respond to questions;
- listen to and comprehend stories;
- develop phonological awareness skills;
- read and comprehend grade-appropriate text;
- engage in emergent and creative writing, e.g. draw, describe pictures, use words in sentences, and use writing prompts to create stories.

Specific and detailed lesson plans were created in accordance with these literacy goals.

A collaborative action research programme

Based on the above criteria, a collaborative, action research programme was initiated in 2014 in Allahabad, Uttar Pradesh to inform the development of a literacy and numeracy programme for children in Grades 1–3. This phase built upon a year-long pilot programme that was conducted in Dadri, Uttar Pradesh. The longitudinal project is in its third year of implementation, with the baseline cohort (who began as first-graders) currently in Grade 3. Programme teachers, selected from the children's local community, were trained and paid by Pratham and are active collaborators along with Pratham field staff in developing, revising, and fine-tuning classroom practices. The longitudinal nature of

the study allows us to understand the progression of literacy skills across the early primary years, as well as to better understand the role of sub-skills in the 'learning to read' and comprehension process.

The programme was implemented in 10 rural and 5 urban government schools[19] where the language of instruction is Hindi. Children in Grades 1 and 2 in these schools received five days of literacy and numeracy enrichment activities for four hours, split equally between literacy and numeracy. Children in an additional 15 schools (10 rural and 5 urban schools) received a mitigated version of the programme with only one day of scheduled activities. An additional 10 schools (7 rural and 3 urban) participated as control schools and received no material or training benefits. Note that all children from Grade 3 (excluding control schools) onward receive only one day of enrichment activities facilitated by the Pratham teacher, but continue to have access to reading materials and worksheets on all days. This paper focuses primarily on the literacy component of the programme for the group that received five days of literacy enrichment.

Teacher selection

Programme teachers were selected from the children's communities and had to have a minimum of 12 years of experience in education. During a two-month training programme, they worked in other Pratham programmes where they had the opportunity to observe and assist experienced teachers. The selection of teachers from this group was based on observations of their teaching. Programme supervisors recruited teachers who demonstrated a more engaging, child-oriented approach, as opposed to a didactic approach to teaching.

Study participants and home backgrounds

Two hundred and seventy children were assessed at all four time points. Of this group, the families of 246 were interviewed for the household survey. Seventy-nine per cent of children were girls and 57 per cent resided in rural areas. Ninety-eight per cent of families reported Hindi as their home language. However, this number is likely inflated, as during their work with the children the field staff and teachers realized that many of them spoke regional dialects rather than the standard variety, and parents tended not to distinguish between the spoken and the standard variety.

19. Schools were selected on the basis of enrolment (at a minimum 25 students) and teaching staff willing to commit to a long-term project. The schools also did not participate in any other literacy and numeracy support programmes provided by the government or other NGOs.

A substantial 79 per cent of mothers and a relatively lower 38 per cent of fathers reported never having attended formal school. Seventy-nine per cent of fathers were employed while only 29 per cent of mothers reported working, either outside the house or from home.

Availability of print materials, especially children's reading materials, at home were extremely limited (see *Table 6.1*). A substantial 79 per cent and 81 per cent of families reported never engaging in storybook reading or storytelling, and 43 per cent of the participating children never borrowed any reading materials from school. Cumulatively, all these indices suggest that the study participants truly represent the bottom of the pyramid.

Table 6.1 Types of print materials and their availability at home

Calendar	98%
Newspaper	7%
Religious text	18%
Picture books	2%
Number books	9%
Story books	3%
School books / workbooks	80%

Source: Authors' calculations. Data from the Early Years' Longitudinal, Collaborative, Action Research Programme.

The assessments: design, development, and associated challenges

The key assessment challenges for this project have been developing orthography- and programme-specific assessments that account for children's home languages, and allow for the tracking of learning gains over time to help discriminate along the learning spectrum. A notable challenge in the design of these assessments has been the tension between the scope and length of assessment versus testing time requirements. The assessments also needed to be easy for teachers to administer and score. The assessments were developed during the pilot phase of the project through a collaborative process of field testing. In alignment with the literacy goals, the assessments were designed to capture oral language ability, conceptual understandings of the uses of print, ability to manipulate the sounds of a language (i.e. phonological awareness), decoding ability, reading comprehension, and writing ability. *Table 6.2* lists the 'skill sets' that were assessed at each time point.

Research programme

To develop the assessments of children's receptive and expressive vocabulary knowledge via picture identification and picture naming tasks, the initial pilot work involved listing local vernacular labels for all pictures based on feedback from teachers. A graded scoring scheme was then developed which allowed us to provide partial credit for responses in the vernacular dialect/language so as to arrive at a better estimate of children's conceptual knowledge.

In addition, for the noun definition task, the prompt was presented as a picture rather than a word to circumvent the possibility that the child might fail to relate to the label in the standard language variety and record a non-response even though he or she had a strong conceptual understanding of the noun prompt.

Children's ability to decode was evaluated using the assessment developed for the Annual Status of Education Report (ASER) survey (see www.asercentre.org for details). In accordance with the pace of reading acquisition mandated in state-prescribed textbooks, the ASER reading assessment for first graders focuses on words with simple akshara, and the assessment for second graders includes words with complex akshara, i.e. consonant-vowel ligature pairs that constitute the barakhadi.

All assessments, with the exception of the written assessments, were orally and individually conducted by programme teachers, and required between 20 and 25 minutes to administer. The assessment of emergent writing (Grades 1 and 2) and the assessment of reading comprehension and sentence construction (Grade 2) were group-administered in a separate session. This strategy helped reduce the overall testing time for teachers and ensured that the duration of a single testing session was reasonable for young children. Teachers were trained to administer the assessments in a standardized manner. The scoring by teachers was cross-verified by a team of four Pratham field personnel and, in addition, a random sample of five children were re-assessed by the field staff at each time point. Children participating in the five-day literacy programme were assessed at the beginning and end of Grade 1 and at the beginning and end of Grade 2. Children in the control group were assessed at the beginning of Grade 1 and at the end of Grade 2. In addition to child assessments, programme teachers maintained logs of their classroom practices and observations.

Table 6.2 Map of the literacy skills assessed at each time point

	Baseline: beginning of Grade 1	Midline 1: end of Grade 1	Midline 2: beginning of Grade 2	Endline: end of Grade 2
Concepts about print (6 items)	Y (.86)	Y (.84)	Y (.84)	Y (.88)
Phonological awareness: syllabification (3 items)	Y (.98)[a]	Y (.99)[a]	Y (.98)[a]	Y (.96)[a]
Phonological awareness: syllable blending (3 items)	Y	Y	Y	Y
Phonological awareness: beginning syllable (3 items)	Y	Y	Y	Y
Phonological awareness: final syllable (3 items)	Y	Y	Y	Y
Picture identification (10 items)	Y	Y	N	N
Picture naming (10 items)	Y (.72)[b]	Y (.71)[b]	N	N
Noun definition (1 item)	Y	Y	Y	Y
Emergent writing (1 item)	Y	Y	Y	Y
Reading (decoding; ASER survey assessment)	(a) Akshara (b) Simple words (c) Words with consonant-vowel ligatures (d) Grade 1 level text	(a) Akshara (b) Simple words (c) Words with consonant-vowel ligatures (d) Grade 1 level text	(a) Akshara (b) Simple words (c) Words with consonant-vowel ligatures (d) Grade 1 level text (e) Grade 2 level text	(a) Akshara (b) Simple words (c) Words with consonant-vowel ligatures (d) Grade 1 level text (e) Grade 2 level text
Listening comprehension (oral response) (3 items)	Y (.82)	Y (.90)	Y (.79)	Y (.93)
Listening comprehension (written response) (3 items)	Y (0)[c]	Y (.95)	Y (.94)	Y (.97)
Reading comprehension (2 items)	N	N	Y (.78)	Y (.90)
Sentence construction (2 items)	N	N	Y (.89)	Y (.94)
Reliability estimates for the complete battery	(.79)	(.88)	(.91)	(.92)

Source: Authors' calculations. Data from the Early Years' Longitudinal Collaborative Action Research Programme.
Note: Reliability estimates based on internal consistency are indicated in brackets, where appropriate.
a: The estimate is for the set of phonological awareness tasks. b: The estimate is for the set of vocabulary measures.
c: The 0 estimate is due to substantial floor effects at baseline.

How have these assessments enhanced our understanding of learning?

A data-driven understanding of what works in the classroom is certainly of great value. Equally important are insights gained from engaging teachers in the tracking and monitoring of student progress, as it allows them to make the very important connection between instruction, student learning, and goal-setting. In the ensuing discussion we highlight some of these insights.

Teacher engagement

At baseline, during the sessions that trained teachers to administer the assessments, much resistance was noted towards the assessment of print and phonological awareness concepts. The general reaction of teachers was that a literacy programme should focus on assessing reading and writing skills. To them, concepts about print and phonological awareness added little value to the monitoring process, and increased the time required for testing. Interestingly, as the programme progressed, teachers became attuned to the demonstration of these skills in their students and also began relying on these as indicators to discuss progress and reading outcomes for individual children. Teachers' initial reservations about these skills, as triggered by the assessments, appear to have sensitized them to the relevance of these skills, and have also inspired classroom activities. For instance, syllable manipulation activities in the classroom were more enthusiastically embraced over time. In fact, one programme teacher coupled the syllable manipulation task with getting children to track down the consonant-ligature vowel syllable unit on the barakhadi (consonant and vowel ligature matrix) chart. The connection between the aural and visual representation was perceived to be very useful, and as a consequence this activity was extended to all programme classrooms.

Teachers were encouraged to maintain daily logs that documented their evaluation of the day's activities. Teachers were also encouraged to write about two or three students in their logbook each day as a way to understand their needs or strengths. Thus, a localized, skills-oriented approach to assessment enabled teachers to extend the learnings from assessments to inform decision-making and instructional practices

Home language

Two schools in the five-day literacy enrichment programme were notable for the wide prevalence of the community dialect, *Avadhi*. At baseline, due to the oral and individual administration of the assessments, programme

teachers who were fluent in *Avadhi* were able to identify the gap between children's oral language proficiency in the community language and in the standard school variety. Programme teachers in these two schools then made concerted efforts to sensitize children to the differences in the language of the community and that of the school. For instance, a frequently occurring activity in these classrooms was to divide children into groups and have the first group list words in *Avadhi* and the second group provide their equivalent in the school language. A word wall was maintained that listed these words under 'I say' and 'it can also be called' or 'what the principal calls it' (the school principal in this case being used as a referent for Hindi speakers). Pratham field staff and supervisors noted that it was common for these children to respond to questions by first asking, 'Should I say it in my language or how the principal says it/how it is written in the book?' The assessments, which allowed partial credit for demonstrations of lexical knowledge in the community language, helped alert teachers to their students' language demands and inform their instructional focus. By engaging in activities that valued children's home language, teachers were able to capitalize on their conceptual knowledge to foster vocabulary acquisition in the school language, and also sensitize children to the differences in the contexts of use for the two languages.

Skill set

The sample of skills assessed was based on our review of literacy acquisition research and the research base for akshara languages. An evaluation of the skills in relation to children's ability to decode and to make meaning from the text validates the importance of these skills for the reading and meaning-making process. See the positive and significant correlation estimates listed in *Tables 6.3* and *6.4* with the exception for baseline, where the estimates are attenuated due to strong floor effects. The importance of these skills is further corroborated by multiple regression analyses. These analyses indicate that good comprehenders demonstrate better phonological awareness skills, vocabulary, reading accuracy, and writing ability, which is consistent with findings for the akshara languages (see Vagh, Nag, and Banerji, 2017) as well as for other languages (e.g. English: Snow, Burns, and Griffin, 1998). These skills cumulatively account for a substantial 77 to 83 per cent of the total variation in children's ability to comprehend text (*Table 6.5*). In sum, these findings suggest that the assessments were successful in capturing skills that are related to the fundamental purpose of reading – the ability to make meaning – which validates the stated literacy goals of the programme.

Table 6.3 Simple correlation coefficient estimates of children's reading levels with other literacy skills (n=270)

Reading	Concepts about print	Phonological awareness	Vocabulary	Emergent writing	Spelling/sentence construction
Grade 1 baseline	0.3	0.3	0.4	0.2	0.4
Grade 1 endline	0.5	0.5	0.5	0.4	0.7
Grade 2 baseline	0.5	0.7	---	0.5	0.6
Grade 2 endline	0.6	0.7	---	0.5	0.7

Source: Authors' calculations. Data from the Early Years' Longitudinal Collaborative Action Research Programme.

Table 6.4 Simple correlation coefficient estimates (Pearson's r) of the comprehension measures – listening comprehension (LC) and reading comprehension (RC) – with other literacy skills (n=270)

		Concepts about print	Phonological awareness	Vocabulary	Reading	Emergent writing	Sentence construction
Grade 1 baseline	LC (oral)	0.4	0.4	0.3ns	0.3ns	0.3ns	---
	LC (written)	0.1ns	0.1ns	0.4	0.1ns	0	---
Grade 1 endline	LC (oral)	0.6	0.6	0.5	0.6	0.4	---
	LC (written)	0.4	0.4	0.5	0.8	0.3	---
Grade 2 baseline	LC (oral)	0.4	0.6	---	0.6	0.5	0.5
	LC (written)	0.4	0.6	---	0.8	0.6	0.7
	RC	0.4	0.5	---	0.6	0.4	0.7
Grade 2 endline	LC (oral)	0.6	0.6	---	0.6	0.5	0.5
	LC (written)	0.5	0.6	---	0.7	0.7	0.8
	RC	0.4	0.5	---	0.6	0.7	0.8

Source: Authors' calculations. Data from the Early Years' Longitudinal Collaborative Action Research Programme.

Table 6.5 Regression models predicting comprehension based on concurrent assessments of concepts about print, phonological awareness, reading level, writing ability, and end-of-Grade 1 assessment of vocabulary (n=270)

	Comprehension B (se)	
	Grade 2 baseline	**Grade 2 endline**
Intercept	-1.2 (0.9)	-2.8** (0.9)
Vocabulary[a]	0.1** (0.04)	0.1*** (0.03)
Concepts about print[b,c]	0.8* (0.4)	1.6** (0.6)
Phonological awareness[c]	0.1* (0.06)	0.1** (0.07)
Reading level (ASER)[c]	0.8*** (0.1)	0.5*** (0.1)
Emergent writing[c]	1.0** (0.4)	2.4*** (0.4)
Sentence construction[c]	1.2*** (0.1)	1.3*** (0.1)
R-square (%)	77%	83%

*p<.05, **p<.01, ***p<.001

Source: Authors' calculations. Data from the Early Years' Longitudinal Collaborative Action Research Programme.

Notes: 'Comprehension' is a composite of reading and listening comprehension. a: Vocabulary was assessed at the end of Grade 1. b: Concepts about print is indexed by a single item, word-to-word mapping, as the rest of the items had strong ceiling effects ranging from 91 per cent to 97 per cent. c: These are concurrently assessed measures.

Challenges related to test development and test administration

Familiarity with test formats

The baseline assessment in our study is not a 'true' baseline as it was conducted a month after school commenced. The pilot phase in Dadri highlighted the fact that, for most children, it was their first experience with a formal system of education and for some, it was the first time anyone in their family had ever attended school. Children tended to be hesitant when interacting with the teacher or when participating in planned activities given the unfamiliar setting of a classroom. Hence, the first month of the programme was designated as a warm-up phase with the objective of familiarizing children to the classroom setting and to the teacher. A few activities also mimicked some of the assessment formats (e.g. identifying pictures in books, naming pictures, playing games involving sound manipulations). This was done so that a lack of familiarity with the testing format did not undermine performance. The warm-up phase was conducted in intervention and control schools.

Oral language: Assessments of oral language proficiency pose considerable administrative and scoring challenges. Among the spectrum

of oral language measures, picture identification and picture naming are the easiest to administer, yet the breadth of the domain and the challenges of a multilingual context present substantial challenges. Multilingual children's conceptual knowledge base is most likely spread out across language varieties, which need to be accounted for in order to better understand their word and world knowledge (Snow, 2017). In our work, we have tried to account for children's home languages by providing partial credit for all the possible vernacular variants in use in the community. The limitation of this approach is that it restricts the generalizability of the measure to other contexts where other variants may be in use.

Comprehension

Studies have indicated that not all comprehension measures are created equal. They tend to differ in the extent to which they capture decoding skills versus oral language comprehension skills (e.g. Keenan, Betjemann, and Olson, 2008). In this project, responses to the comprehension measures were captured orally and in written format. The written format comprehension measures, it can reasonably be argued, bring a third factor into play: that is, children's writing ability.

Concepts about print

By the end of Grade 1, most children demonstrated knowledge of five of the six print concepts. Despite high ceiling effects, the print concepts task was administered at all time points, as it served as a good warm-up activity to begin the assessment session and, moreover, we were interested in tracking the progression of the sixth item, children's ability to point word-by-word as the examiner reads the text aloud.

Scalability

The process of conducting assessments, orally and individually, helped programme teachers to better understand the individual needs of children, sensitized them to the relevant sub-skills, and helped tailor their pedagogy and set goals. The teachers in this programme also had a great deal of flexibility to operationalize these aspects, and they also found the assessments to be easy to administer and score. However, the current project is a single-site, small-scale research programme that serves as a testing ground to develop an early years programme. In the current research study, while contextualized assessments have helped enhance our understanding of learning at the BoP, an important consideration is the scalability of the effective features of the programme, including

the assessment module. Currently, the Early Years Programme is being implemented in two formats. The first is in partnership with state governments, where government teachers are trained to implement the programme, such as in Malda district, West Bengal; Kullu district, Himachal Pradesh; several districts in Jharkhand; and a few districts in Uttar Pradesh. The second is implemented directly by Pratham in 11 Indian states.

Conclusion

The uptake of the programme features has been adapted to the local context, language of instruction, and implementation modality. In classrooms where government teachers implement the programme, efforts have been made to integrate the programme with the state-mandated curriculum. As Pratham and government partnerships are scaling up, the team is increasingly involved in adapting the learning from the research and development phase to a large-scale setting, and dealing with the unique challenges of each context. The assessment module has also been modified and adapted across the implementation contexts. An assessment that is individually administered and requires 20–25 minutes may not be feasible for government school teachers given class size, multi-grade classrooms, and existing state-mandated assessments. Concessions have been made in length and breadth of skill sets assessed in consultation with government teachers, who are the ultimate implementers and consumers of test results.

It is important to note that teachers have viewed the assessment data as actionable and relevant not only for setting goals, but also for innovating instructional practices relevant to attaining the desired goals. When assessments are external evaluations, it is possible that teachers believe understanding and reacting to the data is someone else's job. Hence, key strengths of the Early Years Programme are that not only have localized assessments enhanced our understanding of children's learning at the BoP, but also assessments have been dynamically used by teachers to serve as an action-generating tool. An ongoing challenge is to determine how best a condensed version of the assessment can continue to inform classroom instructional practices effectively as well as help us to monitor the effectiveness of the programme at scale.

References

Jenkins, J.R.; Fuchs, L.S.; van den Broek, P.; Espin, C.; Deno, S.L. 2003. 'Sources of individual differences in reading comprehension and reading fluency'. In: *Journal of Educational Psychology, 95.*

Keenan, J.M.; Betjemann, R.S.; Olson, R.K. 2008. 'Reading comprehension tests vary in the skills they assess: differential dependence on decoding and oral comprehension'. In: *Scientific Studies of Reading,* 12, 281–300.

Nag, S. 2007. 'Early reading in Kannada: the pace of acquisition of orthographic knowledge and phonemic awareness'. In: *Journal of Research in Reading, 30*(1), 7–22.

——. 2013. 'Akshara–phonology mappings: the common yet uncommon case of the consonant cluster'. In: *Writing Systems Research, 6*, 105–119.

Nag, S.; Snowling, M.J. 2012. 'Reading in an alphasyllabary: implications for a language-universal theory of learning to read'. In: *Scientific Studies of Reading, 16*(5), 404–423.

Patel, P.G. 2004. *Reading acquisition in India: models of learning and dyslexia*. New Delhi: Sage.

Prahalad, C.K. 2006. *The .fortune at the bottom of the pyramid: eradicating poverty through profits.* Upper Saddle River, N.J: Wharton School Publishing.

Salomon, R. 2000. 'Typological observations on the Indic script group and its relationship to other alphasyllabaries'. In: *Studies in the Linguistic Sciences, 30*(1), 87–104.

Snow, C.E. 2017. 'The role of vocabulary versus knowledge in children's language learning: a fifty-year perspective'. In: *Infancia y Aprendizaje, 40*(1), 1–18.

Snow, C.E.; Burns, M.S.; Griffin, P. (Eds). 1998. *Preventing reading difficulties in young children.* Washington DC: National Academy Press.

Teale, W.H.; Sulzby, E. 1986. 'Introduction: emergent literacy as a perspective for examining how children become writers and readers'. In: W.H. Teale and E. Sulzby (Eds), *Emergent literacy: writing and reading* (pp. vii–xxv). Norwood, N.J.: Ablex.

Vagh, S.B.; Biancarosa, G. 2012. 'Early literacy in Hindi: the role of oral reading fluency'. Poster presentation for the International Association for the Study of Child Language, Montreal, Canada.

Vagh, S.B.; Nag, S.; Banerji, R. 2017. 'India: the policy and practice of early literacy acquisition in the akshara languages'. In: N. Kucirkova, C.E. Snow, V. Grover, and C. McBride (Eds), *The Routledge international handbook of early literacy education: a contemporary guide to literacy teaching and interventions in a global context* (pp. 233–246). New York: Routledge.

Commentary

Penelope Bender

Vagh and Sharma's paper reviews a localized assessment approach that Pratham is implementing in conjunction with a reading programme in Dadri, Uttar Pradesh. The assessment design is localized through attention to characteristics of the language of instruction and children's home languages/dialects, and contextualized through consideration of children's home backgrounds. The assessment approach is designed to strengthen teachers' understanding of the links between instructional goals, instructional activities, and student learning, and teachers' use of effective instructional approaches.

The paper contributes to programmatic literature on early grade reading through analysis of the structure of Hindi and its implications for reading instruction. Many reading programmes funded by the international community since 2011 have been designed from an English-language perspective, using transcribed African languages as languages of instruction. These languages use the Latin alphabet, with additional characters for some languages. Another group of programmes has been designed for Modern Standard Arabic, following a similar approach. The phonemic and phonological awareness components of those programmes are not easily adaptable to the structure of *ashkara* (Indian) languages. The issues of dialect and diglossia have not received much attention in many of these externally funded programmes, despite the instructional challenges they pose. The paper's illumination of these challenges and description of some possible solutions is useful not only for this context, but for programmes in other regions of India and around the world. The other instructional observations in the paper, such as the importance of oral language and emergent writing, could also be reinforced in existing reading programmes.

The paper also contributes to the literature through its description of teachers' reactions to the assessment items and to student results. Frequently, formal assessment data is collected by external actors, who take the data and never return. Teachers and students rarely see assessment results, and most programmatic assessments are designed so that schools and students cannot be identified. While this is done to protect schools and students, it does not allow teachers to connect with the data in the same way as they do when they collect it for their own students, even when they are provided with national or regional

results. Vagh and Sharma describe the ways teachers in the Pratham programme began to engage students in activities designed to foster understanding of concepts of print and phonological awareness, as they saw the importance of these precursor skills for the mastery of higher-level skills.

As the authors mention, the use of a teacher-administered, 20–25 minute, individually administered oral assessment is not practical as a long-term approach. There are existing assessments that are better suited for frequent classroom use. It is also possible to limit the assessment by focusing on content that has been taught, once teachers are aware of the importance of the range of skills for children's reading acquisition. An additional option is adding frequent, very simple 'checking for understanding' prompts. Eventually, as teachers' ownership of the instructional programme increases, simple assessments can be used by teachers, with external enumerators conducting lengthier assessments much less frequently.

Commentary

John Mugo

Vagh and Sharma take us to the bottom of the pyramid in India, and build a strong case for localized learning assessments in the monitoring of Sustainable Development Goal (SDG) 4. The authors take the literacy acquisition battle to Indian communities where 79 per cent of mothers and 38 per cent of fathers never attended school, and where most children had never stepped in school. This context justifies the claim made by the authors, that literacy assessments must be sensitive to the local and contextual needs of language acquisition.

The chapter is based on the *akshara* languages project, which aims to take into consideration children's current level of print and phonological awareness as a way of accelerating learning to read. The authors stretch the theoretical underpinnings to reading acquisition from an initial three-pillar thesis consisting of engagement with print, scaffolding by adults, and direct instruction, to the additional two components of oral language proficiency and the learner's own readiness and motivation to learn.

The evidence presented in this chapter raises three challenges to reading instruction: the extent to which teachers: (1) foster syllabic and phonemic awareness to systematically cover the entire phonology of their language, (2) balance between the oral language dialect used in the community and the written local language learned at school, and (3) utilize assessment to identify and teach children at the BoP. In my assessment, the evidence responds well to all these challenges.

One important lesson provided by the chapter is that teachers' attitudes towards the value of assessment in improving instruction can be transformed. The initial attitude that measurement of phonological awareness lacks value, and that assessments should be limited to reading and writing, transformed to an appreciation and application of measuring print and phonological awareness. This confirms the adage that what gets measured wins attention.

From the evidence presented in this chapter, we are able to unearth two hidden insights for literacy assessment at the BoP. One, assessments should measure not just what is easy to measure (reading and comprehension), but what is important. Second, assessment capacity must be transmitted to teachers. The collaborative effort among Pratham and Indian Government teachers yields an important transition from

'measurement of learning' to 'measurement for learning'. This is made possible through the deliberate teacher involvement and strengthening of the teachers' capacity to assess.

The chapter leaves us with three key lessons for the application of localized learning assessments to track SDG 4. First, localized assessments can inform and become part of national systems. Second, small pilot assessments can evolve into more ambitious measurements at scale. Third, as localized assessments move to scale, adaptations are possible to incrementally embrace the realities of government schools and the challenges of local languages.

Chapter 7
Ethnography and assessment at the bottom of the pyramid

Bryan Maddox

Introduction

This paper will discuss the contribution of ethnographic research and social and cultural anthropology theory and methodology to the assessment of learning at the 'bottom of the pyramid'. The paper is presented in two parts. In the first part the paper will discuss some of the challenges of conducting educational assessment at the BoP, with a particular focus on issues of diversity. It includes a discussion of anthropological themes of 'context', 'difference' and 'localization', and their relevance to anticipating and understanding the challenges of 'cultural fit' related to assessment, and the moments when large-scale, standardized assessment meets 'local', vernacular practice. In other words, it discusses the typical territory of anthropological critique and points out the challenges and limitations of conventional assessment practice. The paper questions the notion of marginalization as an essential fact, and instead considers it as a relationship to, or outcome of, the practices of powerful assessment institutions. The second part of the paper looks at how ethnographic research can provide practical solutions to the challenges of conducting assessment with groups at the BoP, and in particular, how it can help assessment programmes with the particular challenges of assessment within and across diverse cultural settings.

The paper draws upon my experience as an ethnographer of small-scale, informal assessment – of the type that Dan Wagner might describe as 'smaller, quicker, and cheaper' forms of assessment (Wagner, 2011). This perspective has evolved, informed by lengthy periods of ethnographic research in South Asia (Bangladesh and Nepal). The paper is also informed by more recent ethnographic research with large-scale assessment programmes, namely UNESCO-UIS's LAMP (Literacy Assessment and Monitoring Programme) and OECD's PIAAC (Programme for the International Assessment of Adult Competencies). The challenges and solutions that I discuss in this paper are informed

by ethnographic research with those programmes (e.g. Maddox, 2014; 2015a, Maddox and Zumbo, 2017).

The broad theme of assessment at the BoP is inherently appealing in light of the Rawlsian principle that social resources and attention should be directed to those who are least well-off (Rawls, 1971). Until now, those principles have received insufficient attention in the realm of assessment design. Their profound implications (discussed in this volume) are therefore particularly welcome. Nevertheless, I do not make the a priori assumption that expensive and time-consuming, large-scale assessments are *necessarily* unhelpful or counter-productive for groups at the BoP. As other chapters in this volume illustrate, designing and conducting assessment for groups at the BoP is inherently difficult. As a result, it requires significant time and investment to provide quality assessment and to deliver benefits for those who are being assessed.

This chapter takes the view that many of the challenges that we deal with when discussing assessment at the BoP relate to a set of *differences* – whether we are talking about differences of ability, or those of language, geographical location, disability, ethnicity, or religion – and those differences are present within and across countries. This theme highlights a set of questions. To what extent are assessment practices and regimes sufficiently able to recognize and include these kinds of differences, without producing relations of disadvantage (i.e. issues of test fairness and procedures for anticipating and removing sources of test bias)? These issues are already recognized by major testing institutions. We might also consider a 'horizontal' perspective on difference. That is, is there anything particularly distinctive from an assessment perspective about those groups who are somehow at the 'bottom' of a social economic system? The answer to that question is by no means straightforward.

I am wary about considering a certain *set of differences* as necessarily *equivalent* to marginalization or educational disadvantage. For example, a person with a disability or impairment may or may not be disadvantaged or marginalized in terms of their social economic position or their educational experience. Similarly, the fact that a person speaks a 'local' or 'minority' language does not necessarily mean that their participation in a large-scale assessment is illegitimate or invalid. In both cases, we might argue that the primary assessment task is to provide inclusive and robust practices that enable the person's skills and abilities to be properly assessed and recognized. The commitment to the quality and reliability of assessment practices is especially important when we

consider Messick's concerns about the responsibilities of assessment programmes to consider the intended or unintended consequences of assessment design (Messick, 1998; Hubley and Zumbo, 2011; Kane, 2016).

Secondly, I am hesitant to adopt the language of the periphery that implies there is some place – a centre – where issues of difference are no longer important. I doubt that such a place exists. As Bruno Latour (1996) has argued, we must be careful about how we use notions of the 'local' as distinct from the 'global'. In practice, what we have is a set of interconnected and heterogeneous local contexts – some of which play an active part in the design of assessment programmes and in the interpretation of their results.

It is not clear how a hierarchical notion of a 'pyramid' connects with the kind of flat ontology that such heterogeneity implies (Moll, 2002). However, it is clear that some local contexts and groups are more influential in shaping the content and aims of assessment programmes than others. That suggests that there is a need to understand the implicit or tacit cultural knowledge and logic that informs assessment design, assessment constructs, and test items.

Assessing dominant and vernacular practices

My research has explored these questions (such as test bias and cultural fit) in relation to the question of how effectively tests travel – that is, how they are received within and across different cultural contexts and settings. These themes of 'top-down' meets 'bottom-up' vernacular perspectives have been discussed widely in ethnographic research on language and literacy, and highlight important issues of power, diversity, and representation (e.g. Shohamy, 2004; McNamara and Roever, 2006; Blommaert, 2008; Street, 2011; Barton and Hamilton, 2012; Hornberger and Johnson, 2011).

Developers of standardized tests generally have an aversion to designing items that might only be meaningful within a limited context – and this is particularly evident in large-scale assessments when they must be designed to be reliable across diverse socio-cultural and linguistic settings. Furthermore, introducing 'local' context into large-scale assessment (e.g. as in the contextual fabrication of assessment items) is not an easy practice. As my ethnographic research in Mongolia showed (Maddox, 2014), 'going local' in assessment design (or for that matter going peripheral or marginal) can introduce unexpected sources

of bias and misfit, even when the 'local' test item content may be thought of as the most appropriate and familiar (discussed below).

This has parallels in literacy research, such as in the distinction between 'dominant' and 'vernacular' literacies (Barton and Hamilton, 2012). The vernacular is associated with unruly, small-scale local practices, which are idiosyncratic in nature, and which contrast with the dominant literacies of powerful institutions. The challenge of assessing 'competence' in the movement between dominant and vernacular forms has been discussed in the ethnographic work of Blommaert (2008). As that literature demonstrates, while people may be more familiar with local or vernacular practice, they may also have to engage from time to time with more dominant, standardized, and powerful forms of literacy, such as in encounters with the official documentation of state institutions.

This issue of vernacular versus dominant literacies is particularly important for assessment at the BoP, because low levels of literacy or numeracy ability in contexts of poor-quality schooling may well be associated with (a) vernacular, localized practices that are not understood or captured within large-scale assessment; and (b) complex forms of engagement with dominant practices (such as collective and mediated encounters). There are issues of power to be considered in terms of *whose* priorities and *whose* literacies, languages, and scripts are being assessed and why (Street, 2011).

This raises a question about how people's existing abilities might be understood and assessed effectively. Dominant social groups (such as those who design assessment programmes) may not be sufficiently competent and familiar with the forms of literacy that operate in local contexts. The implication is that low levels of ability are not simply difficult to capture accurately in large-scale standardized testing – they may also be mistaken as an *absence* of skills if assessment scales are not adequately populated with appropriate test items, or sufficiently nuanced to capture vernacular practices at low levels of ability.

Globalizing assessment

My research on the UNESCO LAMP assessment in Mongolia identified a number of challenges with 'globalized' assessments, including problems of cultural misfit associated with 'not enough' context in the assessment items, so that test items appeared alien in the assessment setting (Maddox, 2014, 2015*a*, 2015*b*; Maddox *et al.*, 2015). Ironically,

we also saw examples of 'too much' context, in which the respondents drew too heavily on their local knowledge and expertise and produced incorrect answers. For example, nomadic camel herders were not able to provide the correct answers for test items related to camels because of a discrepancy between their expert local knowledge of camels and the codified information on camels contained in the test items (Maddox, 2014, 2015*b*).

A second concern regarding the notion of the 'local' and assessment at the bottom of the pyramid relates to 'universal' cognitive constructs. I was often taken aback when observing Mongolian nomads successfully completing test items that appeared to be entirely alien to the Mongolian context and culture. In these cases, it appeared that respondents were somehow able to bypass the contextual fabrication of test items to discern the underlying nature of the task. In those cases 'local knowledge' seemed to imply 'deep' cognitive abilities. There are parallel debates in fields such as reading research on the particularity and transferability of cognitive skills. This suggests that diverse groups (or if you like, groups at the BoP) may indeed be able to participate in large-scale, standardized assessment programmes. Being 'local' does not necessarily mean that people cannot engage in the global.

To conclude the first part of this paper: while there are reasons to question the efficacy of large-scale, standardized assessments of learning at the 'bottom of the pyramid', there are also merits to accepting and overcoming the methodological challenges of such an inclusive approach. Consequential questions remain about how (and which) groups benefit from large-scale assessments.

In response to the expansion of international large-scale assessments (ILSAs), there are those who would like to reject large-scale assessment altogether, who express a desire to keep things simple. They want those who are impacted by assessment to benefit from and control its consequences (e.g. see Thompson, 2015). These types of responses seem to imply a concern about the momentum of technological change (the so-called 'assessment juggernaut') and the anxiety about new technologies once expressed by the philosopher Martin Heidegger (1977) – the sense that technological change is being driven by a force somehow beyond human control.

I do not subscribe to such views. However, the movements against large-scale testing do raise legitimate concerns. As Zumbo (2014) has argued, as the infrastructure of large-scale assessment has been

developed in North America and Europe (including the 'standards' and guidelines on ethics and quality) we need to rigorously think through their applicability in other contexts. With respect to assessment at the BoP, we may want to go one step further and consider the tension that may exist between the interests of national elites in low-income countries and those low-educated populations who are supposed to benefit from assessment programmes.

Ethnography and assessment at the BoP

Having raised a lengthy set of complex and perhaps unanswerable questions, in the next part of this paper I hope to provide some solutions in terms of how ethnographic and cultural-social and political perspectives can contribute to assessment research and design.

This section is informed by an ethnographic orientation as to what we can learn from observing 'real-life' assessment in practice. That is, it is informed by a perspective that is quite different from that of observing assessment in the laboratory. This ethnographic perspective relates to what Bruno Zumbo and I have described as assessment 'in-vivo' (Maddox and Zumbo, 2017). It considers assessment practice as it is situated in the messy and noisy contexts in which assessment takes place. In terms of our concern with assessment of learning at the BoP, this perspective implies the adoption of new practices and techniques, such as ethnography, that seek out cultural diversity and the peculiarities of context, not as polluting influences, but as virtuous aspects of assessment design.

Secondly, as tests bump up against the realities of everyday life, they produce important sources of information about the meanings of assessment, and associated content on respondent 'stance' and 'affect' (Maddox and Zumbo, 2017). By using micro-ethnographic and video-ethnographic methods to observe testing situations, we can better understand this process – for example, capturing 'process data' in the talk that takes place during assessment events, and people's facial expressions and body language (Maddox, 2015*a*; Maddox and Zumbo, 2017).

To illustrate this argument, the chapter concludes by presenting some examples drawn from ethnographic observation during the OECD PIAAC assessment in Slovenia. The examples deliberately focus on assessment practices that might be considered characteristic of the BoP As I attempted to identify these cases for this paper from

my field-notes, it is clear that for a number of reasons these individuals may indeed be regarded as marginal – not necessarily in terms of their own worldview, but in how they are integrated into assessment practice. I was particularly interested to observe people who had very low levels of skills in literacy and numeracy. In observing these individuals it was often fairly evident that low levels of skill were associated with poor socio-economic status. These individuals were also 'marginal' in terms of the assessment design. They include people who are not sufficiently competent at literacy, numeracy, or computing to participate in the computer-based assessment. It is clear that for these people completing the assessment task was challenging – producing levels of anxiety that were not typical of the wider study. This reminded me of the anxiety and fear associated with literacy and numeracy that I have witnessed in my research and teaching in the UK and South Asia.

To some extent, these groups at the BoP are distinctive in terms of their educational characteristics, and they challenge assessment programmes that are concerned with inclusion to respond appropriately. In the UNESCO LAMP and OECD PIAAC assessments, there are mechanisms that enable people with low levels of skill to participate. However, it is not always clear whether the assessment design and administration is adequately adapted and sensitive to the vulnerabilities of these groups, such as the need to avoid public shame that may be associated with taking the assessment.

The first example I want to discuss involves a middle-aged man with very low levels of ability who took the PIAAC paper-based assessment in Slovenia. He was not willing to have the assessment event recorded, but did consent to my observing it. The man lived in a partially completed house, with empty beer cans strewn about. He was clearly nervous while taking the test, smoking cigarettes as he did so. Watching him complete the test, it was evident not only that he struggled with literacy and numeracy tasks, but also that those tasks were associated with high levels of anxiety. Before and after the assessment, he talked about his life and about his attempts to become educated despite a difficult childhood. I considered it an honour to have been invited to observe his assessment. In my observations of the PIAAC assessment and the Mongolian LAMP assessment (both household-based adult skills assessments), I was impressed by the commitment of the testing programmes to include people with very low levels of skill in their randomized selection of respondents. That strikes me as being an important characteristic of inclusive assessment, and one that attempts to capture and represent

individuals at the BoP. Even so , such inclusiveness places great demands on the programmes – for example, how well does the assessment content and design match individual needs? These are difficult questions that I believe ethnographic research can help answer.

The second example (the transcript below) is from observing a 'low-skilled' woman as she completed the paper-based option of the Slovenian PIAAC assessment. In this assessment, two things were clear. First, the respondent struggled to answer many of the test items. Second, the interviewer demonstrated considerable skill and commitment towards helping the respondent complete the test. The transcript captures this process and highlights the active role of the interviewer in the assessment process. As I have argued elsewhere, this 'interviewer effect' is one of professionalism and inclusion rather than a source of bias (Maddox, 2017). That is, the interviewer was performing the role as it was set out in the training and in the interviewer manual.

Transcript 1. 'Write the answer below the line'

Interviewer (I):	Write the answer here, below the line.
Respondent (R):	Can I copy it here?
I:	Yes, yes, of course, of course. That's it, and move on to the next exercise. That means here is a question, here you will read and find the answer. If you've done, you can move on. Here is another thing, the fifth question.
R:	I see.
I:	That's it … and now the sixth question.
R:	Hm. How much is the half of this?
I:	Look, I'll give you the calculator. Here. [passes the calculator]
R:	Hm, and if I can't use it … of what use can it be?
I:	No? I see, good! Then I'll give you this sheet of paper! [hands her some paper]
R:	No, it won't do.
I:	No, good.
R:	No, I can't deal with percentages …
I:	But you can round up …
R:	Yeah. 25 … 22 and a half.
I:	Well, let's move on.
R:	But I've messed up everything.
I:	It's okay.

The transcript above shows how much care the interviewer takes in helping the respondent to complete the test. An ethnographic point here is that 'talk' during tests provides very good evidence about the quality of such interaction. In my research, I also use video ethnographic

techniques, and these capture body language and facial expressions to add a further dimension. The spoken interaction was a central and persistent feature of the assessment event.

Transcript 2. 'This makes no sense to me'

R: This makes no sense to me. I don't understand a thing.
I: If you don't understand, just move on, and then you'll …
R: … I will read some more, alright, but …
I: Yeah, yeah, of course, of course!
R: Damn!
I: You will find some things a bit more difficult, some will be a bit easier, so that … if you're interested, yes, read it and try.
R: Ah, this …
I: Yes, it's quite difficult, it's not that simple!
R: Bollocks! I don't even know where to start reading! Where could this one be? I don't know where to look for it.
I: Yes, it is quite difficult. It you can make it, excellent, if not, you can move on.

In the second transcript, you can see that the respondent is really struggling with the demands of the tasks, and making sense of the requirements. She is working very hard. The interviewer provides guidance and encouragement throughout so that the respondent remains engaged in the task. As an example of standardized assessment practice, it is hard to fault the interviewer in terms of her professionalism and commitment to the task. The assessment continues in this style for around 45 minutes, with the interviewer providing ongoing encouragement. At times it seems that the respondent might give up, but each time she is gently encouraged to continue. That encouragement is clear in the transcript below.

Transcript 3.'Whoever invented this, I give him F minus, or less'

I: OK. Let's do this one here then, the ninth question, for which, I believe, you've already found the answer.
R: This one then? Twelve fifths …
I: Yes, that's it.
R: Ay… Bollocks! Do I have to write all this?
I: Yes.
R: Mmm, shoot!
I: So that it includes both reading and writing skills.
R: I'll do the half of it, must I do it all?
I: Good.
R: Bollocks!

R: And I have to write this too? Ah!
I: Cool. Job well done!
R: Whoever invented this, I give him F minus, or less!
I: Really?
R: It makes no sense whatsoever.

This final transcript shows that together, the interviewer and respondent have gotten to the end of the assessment task. Although the interviewer has been active throughout, she did not provide the answers, but instead provided ongoing encouragement and guidance about how to complete the task. Here, the respondent also gives some humorous feedback to the test designers that has a serious message. That is, these types of ethnographic accounts provide an excellent source of qualitative data and feedback about how respondents at the BoP engage with the assessment. The respondent's reaction asks us to consider how well the assessment is designed to suit the abilities of people at the lower end of the ability spectrum (see *Chapter 3*).

I am glad that these kinds of large-scale assessments include respondents with low levels of skill. After all, their inclusion is an essential requirement of such studies. However, the number of expletives in the transcript above suggests that these kinds of large-scale assessments could do more work to adapt their design to fit the needs of groups at the bottom of the educational pyramid. Ethnographic accounts of 'interaction' (McNamara, 1997; Maddox, 2015*a*) in testing situations provide some important 'missing data' on assessment response processes that are not captured by large-scale psychometric analysis or by computer-generated log files. That kind of testimony provides important feedback on how tests are received and potential sources of 'misfit'.

As this paper has argued, designing inclusive and effective (quality) assessments for those at the BoP means dealing with difficult and complex subjects. They must account not only for different levels of ability (if we consider that a trait that the assessment seeks to measure), but also associated types of diversity – of people, ecology, and context – which large-scale assessment programmes are not always entirely sure how to deal with. In Messick's work, 'context' was sometimes regarded as a source of measurement error, as a kind of polluting influence that had to be controlled (Messick, 1989; McNamara and Roever, 2006; Maddox, 2015*a*). The task of dealing with 'construct irrelevant' sources of variation is usually given to committees of 'content specialists', whose role often appears to be brushing problematic examples of cultural misfit under the carpet, or using sophisticated psychometric analyses to make

them go away. An ecological perspective on assessment instead views the interaction between the test, the respondents, and the context as a central concern in making sense of assessment performance and data. That is, it is not so much a problem to be managed away, but a valuable source of ecological information that can be used to improve the quality of assessment practice and data.

Conclusions

We can see from the arguments in this chapter that ethnographic research is not only a valid mode of research, but also a necessary one if assessment at the BoP is to be both reliable and effective. This conclusion is relevant whether we are dealing with the 'smaller, quicker, cheaper' forms of assessment that Wagner and colleagues have advocated, or the more complex, expensive, and larger-scale assessments that are increasingly promoted by the OECD and international organizations.

Ethnographic observations contribute to a wider set of 'process data' techniques that generate ecologically situated – or 'in-vivo' – field-based insights into assessment response processes and quality (Maddox and Zumbo, 2017). Those include detailed accounts of assessment practices and the kinds of interactions that are illustrated in the transcripts in this chapter. As I have shown, ethnographic observation can capture ecologically sensitive data (in the form of ethnographic description and verbal transcripts) about assessment practices that are not captured by the more established techniques of validity practice (e.g. computer-generated log files). As such, ethnography has an important role to play in improving the quality of assessment at the BoP – however we choose to understand that idea.

References

Barton, D.; Hamilton, M. 2012. *Local literacies: reading and writing in one community.* London: Routledge.

Blommaert, J. 2008. *Grassroots literacies: writing, identity and voice in Central Africa*. London and New York: Routledge.

Heidegger, M. 1977. *The question concerning technology.* New York: Harper & Row.

Hornberger, N.H.; Johnson, D.C. 2011. 'The ethnography of language policy'. In: T. L. McCarty (Ed.), *Ethnography and language policy* (pp. 273–289). New York: Routledge.

Hubley, A.M.; Zumbo, B.D. 2011. 'Validity and the consequences of test interpretation and use'. In: *Social Indicators Research,* 103, 219–230.

Kane, M.T. 2016. 'Explicating validity'. In: *Assessment in Education, 23*(2), 198–211.

Latour, B. 1996. 'On inter-objectivity'. In: *Mind, Culture and Society*, 3, 228–245.

Maddox, B. 2014. 'Globalising assessment: an ethnography of literacy assessment, camels and fast food in the Mongolian Gobi'. In: *Comparative Education, 50*(4), 474–489.

——. 2015*a*. 'The neglected situation: assessment performance and interaction in context'. In: *Assessment in Education, 22*(4), 427–443.

——. 2015*b*. 'Inside the assessment machine: the life and times of a test item'. In: M. Hamilton, B. Maddox, and C. Addey (Eds), *Literacy as numbers: researching the politics and practices of international literacy assessment*. Cambridge: Cambridge University Press.

——. 2017. 'Talk and gesture as process data'. In: *Measurement: Interdisciplinary Research and Perspectives, 15* (3–4), 113–127. ISSN 1536-6367.

Maddox, B.; Zumbo, B.D. 2017. 'Observing testing situations: validation as jazz'. In: B.D Zumbo and A.M. Hubley (Eds), *Understanding and investigating response processes in validation research* (pp. 179–192). New York: Springer.

Maddox, B.; Zumbo, B.D.; Tay-Lim, B.S-H.; Qu, D. 2015. 'An anthropologist among the psychometricians: assessment events, ethnography and DIF in the Mongolian Gobi'. In: *International Journal of Testing, 14*(2), 291–309.

McNamara, T. 1997. '"Interaction" in second language performance assessment: whose performance?' In: *Applied Linguistics, 18*, 446–466.

McNamara, T.; Roever, C. 2006. *Language testing: the social dimension.* Malden, Mass. and Oxford: Blackwell.

Messick, S. 1989. 'Validity'. In: R. L. Linn (Ed.), *Educational measurement*, 3rd edn (pp. 13–103). New York: American Educational Council and Macmillan.

——. 1998. 'Test validity: a matter of consequences'. In: *Social Indicators Research, 45*, 35–44.

Moll, A. 2002. *The body multiple: ontology in medical practice*. Durham, N.C.: Duke University Press.

Rawls, J. 1971. *A theory of justice.* Cambridge, Mass. and London: Harvard University Press.

Shohamy, E. 2004. 'Assessment in multicultural societies: applying democratic principles and practices to language testing'. In: B. Norton and K. Toohey (Eds), *Critical pedagogies and language learning* (pp. 72–92). Cambridge: Cambridge University Press.

Street, B.V. 2011. 'Literacy inequalities in theory and practice: the power to name and define'. In: *International Journal of Educational Development, 31*, 580–586.

Thompson, T.; Miles, R.2015. '"I Had No Idea!" A Snapshot of Science Teacher Perceptions of Student Performance on State, National and International Assessments.' *National Teacher Education Journal, 8,* 61–66.

Wagner, D.A. 2011. *Smaller, quicker, cheaper: improving learning assessments in developing countries.* Paris/Washington: UNESCO-IIEP and FTI-Global Partnership for Education.

Zumbo, B.D. 2014. 'What role does, and should the Test Standards play outside the United States of America?' In: *Educational Measurement,* 33(4), 31–33.

Zumbo, B.D.; Hubley, A.M. 2016. 'Bringing consequences and side effects of testing and assessment to the foreground'. In: *Assessment in Education, 23*(2), 299–303.

Commentary

Nancy H. Hornberger

Reading Maddox's essay, I was struck by parallels with the ethnography of language policy which looks across layers of language policy development and implementation, shedding light on how top-down and bottom-up policy interact in particular contexts, and often uncovering indistinct voices, covert motivations, or unintended policy consequences (Hornberger and Johnson, 2007). Following are a few points of connection.

Ethnography as a way of shedding light on cultural bias and misfit in particular contexts

Maddox provides ethnographic examples from Mongolia of how tests do not travel well across cultural contexts, introducing unexpected sources of misfit even where locally appropriate test items are included. The examples are powerful, but unsurprising, since standardized tests are designed with the goal of yielding one-size-fits-all results and have been criticized for lack of local or cultural appropriateness. Indeed, Hymes (1980) made this point when he proposed ethnographic monitoring in prescient anticipation of what would come to be an intractable problem in US bilingual education programme evaluation.

Ethnography as a way of uncovering unintended consequences

In discussing consequential questions about how (and which) groups benefit from assessment, Maddox provides examples from Slovenia where mechanisms enabled people with low-level literacy and numeracy skills to participate. Test interviewers offered guidance, support, and encouragement for respondents to complete the test when they were struggling – and I would argue even distressed. Maddox rightly sees an ecological perspective on interactions between text, respondents, and context as central in making sense of assessment performance and data. The question for me, though, is: to what end?

Is this inclusion directed towards any beneficial consequence for the respondent? While I agree there may be value in including populations at the BoP in large-scale assessments in order to render an overall picture, there is no real argument pointing to benefits for the test-taker who provides data. An ecological perspective includes not only text, respondents, and context, but also the local ecology of language inequality (Hornberger, 2002). Ethnographic research on assessment as language policy joins critical

language-testing research in uncovering troubling consequences of tests for test-takers, bringing attention to the ways in which language tests are not neutral. Rather, they are vehicles for perpetuating the often hidden agendas of those in power by failing to recognize or validate minoritized group members' previous language or content knowledge (Shohamy, 2006).

Final thoughts

Assessment design has a long way to go to offer equitable and responsible assessment for all. It may not even be possible, given the origins and functions of assessment programmes. When I presented an early ethnography of language policy many years ago, I argued that highland Andean Quechua children taught through the medium of Quechua thrived in ways contrasting sharply with traditional Spanish-only medium schooling (Hornberger, 1987). Asked whether I had tested the children to support my findings, I replied I saw no point – and indeed much potential for harm – in doing so, since available tests in Spanish would be not only unintelligible, but also intimidating for the children and therefore yield meaningless results as to what they knew and could do.

In the years since, I have yet to see a body of convincing evidence to the contrary, for any indigenous or minoritized learner population anywhere. I admire Maddox's sensitive and nuanced attention to local context and his commitment to provide quality assessment for all, yet I remain sceptical as to *how* (small- or large-scale) summative assessments as currently envisioned can, in his words, 'produce benefits for those who are being assessed'.

References

Hornberger, N. H. 1987. 'Bilingual education success, but policy failure'. In: *Language in Society, 16*(2), 205–226.

——. 2002. 'Multilingual language policies and the continua of biliteracy: an ecological approach'. In: *Language Policy, 1*(1), 27–51.

Hornberger, N.H.; Johnson, D.C. 2007. 'Slicing the onion ethnographically: layers and spaces in multilingual language education policy and practice'. In: *TESOL Quarterly, 41*(3), 509–532.

Hymes, D.H. 1980. *Language in education: ethnolinguistic essays*. Washington DC: Center for Applied Linguistics.

Shohamy, E. 2006. *Language policy: hidden agendas and new approaches*. London: Routledge.

Commentary

Ameena Ghaffar-Kucher

What does assessment, and therefore learning, look like in the varied contexts at the 'bottom of the pyramid'? As Maddox's rich ethnographic work illustrates, there is no straightforward answer. This is why ethnographic methods matter – they help uncover the aspects of learning (and perhaps unlearning) that we overlook because of our a priori understandings of how learning occurs in contexts that fall outside of the dominant perspective. To be clear, I am not arguing that we simply need to tailor assessments to particular contexts – though we certainly could do more of that. The issue is more fundamental. As Maddox points out, the infrastructure of large-scale assessment has been developed in North America and Europe; consequently, it legitimizes particular forms of knowledge and science that may not be applicable to the heterogeneous BoP (Janson and Paraskeva, 2015). Drawing on the work of de Sousa Santos, João Paraskeva argues that we need a southern epistemology to understand the diverse forms of 'knowledge, cultures and cosmologies in response to different forms of oppression that enact the coloniality of knowledge and power' (de Sousa Santos in Paraskeva, 2014: 84). A tremendous epistemological diversity exists and needs to be made visible – ethnographic work is essential to this endeavour.

This may be the best way to resolve the tension between those who assume that assessment evidence – particularly reading scores – is equivalent to evidence of learning, and those who are pushing us to broaden what we include under the umbrella of learning, beyond reading and numeracy. In fact, as other contributors to this volume have noted, we need to understand marginalized populations and what learning means to them. Helen Boyle's (2004) ethnographic work on Qur'anic schooling is useful to exemplify this point. Contemporary early Qur'anic education focuses on Qur'anic memorization. Though this appears to be rote memorization, Boyle argues that, in fact, Qur'anic memorization is a process of embodying the divine – the words of God – and is actually far more learner-oriented and meaningful a process than it is credited as being. The purpose of contemporary Qur'anic schooling, with Qur'anic memorization at its core, is to develop spirituality, discipline, and a moral compass, as well as a foundation for later public education. Yet by most international standards of assessment in these early years, we would question whether these children are actually learning. Thus we come back to the question: What does learning mean for different groups of people?

Not only do we need to expand our understanding of what learning is in the ecology of assessment, we must also include discussions on teaching and pedagogy and how that informs and is informed by assessment, and on the political and institutional questions and issues of representation and power that Maddox raises. My concern is that with our focus on assessment, particularly large-scale assessment, but even the smaller-cheaper-quicker variety (Wagner, 2011), the complexity of teaching and learning is reduced to quantifiable data. Subsequently, we are increasingly losing our focus on the day-to-day practices of teaching and learning and on the child as the centre of all our efforts.

References

Boyle, H. 2004. *Quranic schools: agents of preservation and change*. New York and London: Routledge Falmer.

Janson, E.; Paraskeva, J. 2015. 'Curriculum counterstrokes and strokes: swimming in nonexistent epistemological rivers'. In: *Policy Futures in Education, 13*(8), 949–967.

Paraskeva, J. 2014. *Conflicts in curriculum theory: challenging hegemonic epistimicides*. New York: Palgrave.

Wagner, D.A. 2011. *Smaller, quicker, cheaper: improving learning assessments in developing countries*. Paris and Washington: UNESCO-IIEP and FTI-Global Partnership for Education.

Chapter 8
Early learning and development at the bottom of the pyramid

Amy Jo Dowd and *Lauren Pisani*

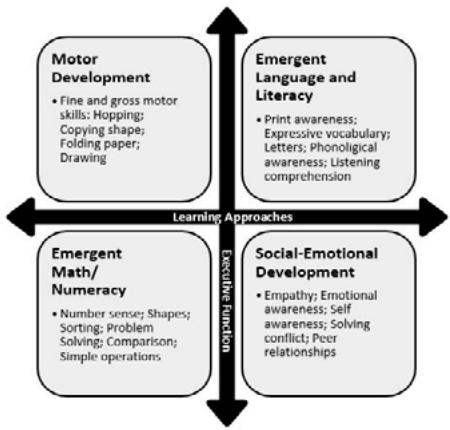
Introduction

While Sustainable Development Goal 4 (SDG 4, Target 4.2) states that every child should have the right to high-quality early childhood/pre-primary education, many children across the world begin their lives without appropriate care and early stimulation. Poverty, stunting, and lack of learning materials and interactions mean that nearly half of 3- and 4-year-old children in low- and middle-income countries (LMICs) are not reaching basic developmental milestones with respect to cognitive, socio-emotional, or physical development, which hinders their potential (McCoy *et al.*, 2016). The evidence tells us that young children learn through relationships with caring, responsive adults; that early positive experiences build strong brain connections (Harvard University Center for the Developing Child, 2016); that warm, responsive, and stimulating caregiving can effectively promote development even in the presence of poverty and malnutrition (Walker *et al.*, 2005; Nores and Barnett, 2010); and that things as simple as talk influence toddler vocabulary (Hart and Risley, 1995) which in turn predicts early grade reading (Heckman and Masterov, 2004). Children's exposure to care and stimulation, play, and positive interactions in the early years determines many key outcomes throughout their lives.

This paper will use data from the International Development and Early Learning Assessment (IDELA) to explore the skills of young children at the BoP (i.e. the most marginalized and disadvantaged). The IDELA tool is a holistic, rigorous, open source, direct child assessment for 3–6-year-old children that is easily adapted and used in different national and cultural contexts. It measures motor development, emergent language and literacy, emergent numeracy, and social-emotional development, with optional questions related to executive function and approaches to learning (see *Figure 8.1* for skill detail by domain).

Save the Children began developing IDELA in 2011 and released it for public use in 2014 (Pisani, Bosirova, and Dowd, 2015). Since then, Save the Children and over two dozen partner organizations have used IDELA for evaluations in 35 countries. IDELA is the focus of ongoing psychometric analyses in partnership with NYU Global TIES for Children.[20] In this paper, we use IDELA data to explore limitations with unidimensional national averages, the multifaceted nature of learning and learning environments, and to offer evidence and examples of effective interventions at the BoP. We will then speak to challenges and limitations related to multidimensionality and comparability.[21]

Figure 8.1 International Development and Early Learning Assessment (IDELA) domains and skills



Source: Pisani, Borisova, and Dowd, 2015.

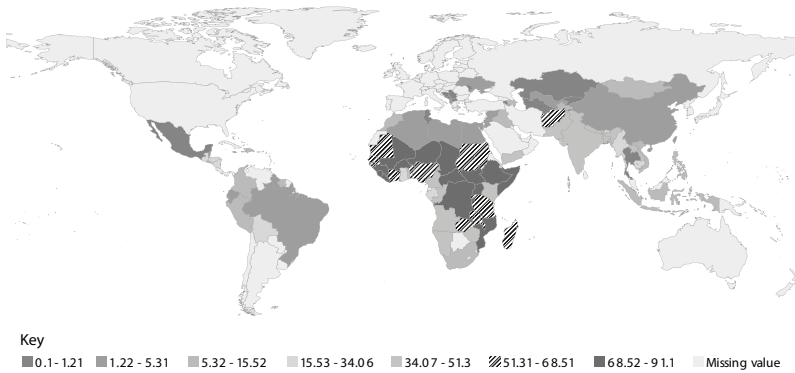
The trouble with national early learning and development averages

Save the Children and many non-governmental organization (NGO) partners working to promote early learning are often situated quite literally at the BoP. The countries shown in *Figure 8.2* with the highest percentages of their populations living in multidimensional poverty are the ones in which Save the Children and its partners mainly work.

20. For detail on NYU Global TIES, see: http://steinhardt.nyu.edu/global-ties.

21. Note that the choice to focus on ages 3–6 years using IDELA in this paper is based on proximity and the extent of our experience. Most of the points made also hold for early learning and development at the BoP for children aged 0–3 years for which we use the CREDI (Caregiver Reported Early Childhood Development Index) tool: https://sites.sph.harvard.edu/credi/.

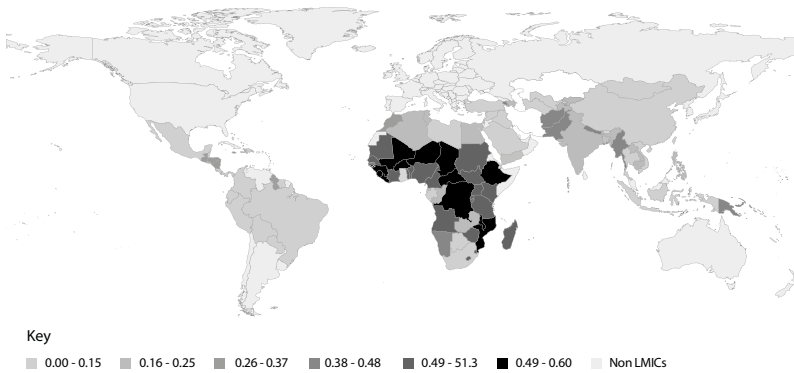
Figure 8.2 Percentage of population living in multidimensional poverty



Source: Oxford Human Poverty and Development Initiative, 2016.

And that map is strikingly similar to the map of estimated low Early Childhood Development Index (ECDI) scores (*Figure 8.3*).

Figure 8.3 Estimated proportion of children with low development per the ECDI by country

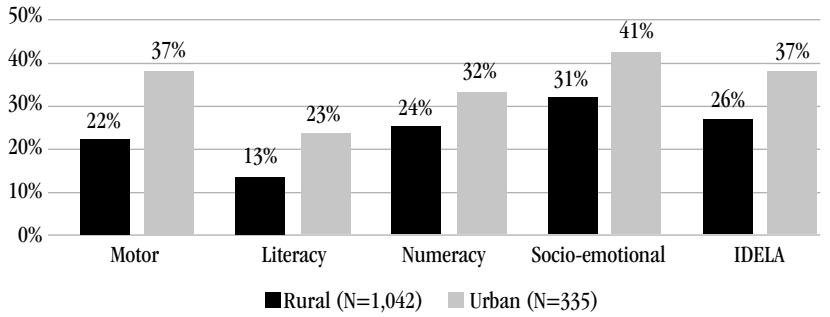


Source: McCoy *et al*., 2016.

Where IDELA has been used on a national scale, we see stark urban–rural divides that cast doubt on the broad utility of national averages for local decision-making. For, although poverty exists everywhere in the countries with the darkest shades in *Figures 8.1* and *8.2,* rural families and their children often face unique challenges, as remoteness often

limits access to public nutrition, health, and early childhood care and development (ECCD) services. A national IDELA study in Bhutan shows rural children lagging approximately 50 per cent behind their urban peers in nearly every domain of the IDELA assessment (*Figure 8.4*).

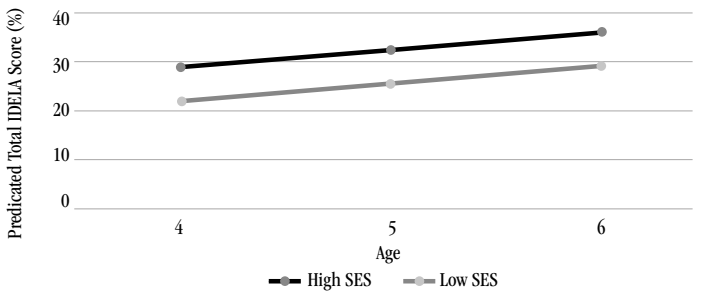
Figure 8.4 Bhutan IDELA results among children in rural and urban areas



Source: Save the Children, 2016.

Further, even in smaller project site samples within rural impoverished areas, we find a range of poverty and its influence on early learning and development. For example, in the Afar Region of Ethiopia, a 2016 IDELA study found that while all children displayed low learning and development, children from low socio-economic status (SES) families are years behind their more affluent peers (*Figure 8.5*).

Figure 8.5 Afar Region, Ethiopia: predicted total IDELA score by SES (n=120)



Source: Save the Children, 2017.

In Afar, a 6-year-old child from a low SES family was predicted to have the same developmental level as a 4-year-old child from a high SES family. Assessments need to capture variation like this at the local level in order to inform programming and advocacy. The magnitudes of these gaps and the low overall level of the scores illustrate the necessity and urgency of finding ways to support the development of economically disadvantaged children.

Looking beyond national averages also highlights that there are often hidden communities of disadvantaged children within urban areas and also within higher-income countries. Partners like the Global Fairness Initiative in Nepal,[22] RISE in Egypt,[23] and the Federal University of São Paulo are using IDELA to investigate marginalized children's learning in urban areas. Also, through collaborations with Open Society Foundation[24] and International Step by Step Association,[25] as well as the World Bank and the Roma Education Fund,[26] organizations across the Balkan region are able to take a closer look at the early learning opportunities provided to Roma children in the region, and work towards providing the high-quality services these children and families desperately need.

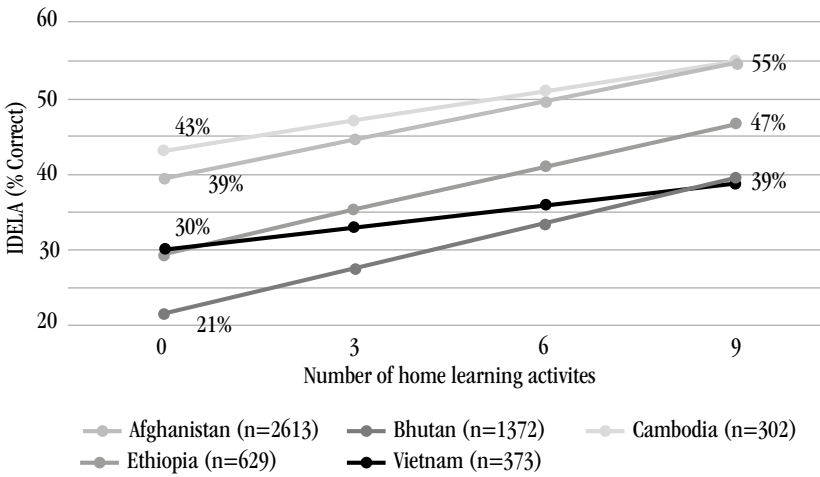
The multifaceted nature of learning and learning environments

Both the nature of learning and the contexts in which learning occurs have many dimensions that require attention in measurement. A key measurement principle of IDELA is that of the multidimensionality of children's early learning and development. Although some stakeholders may view ECCD narrowly as promoting foundational skills such as reading and math for primary schooling (Patrinos, 2017), we advocate directly measuring at least language and literacy, emergent numeracy, and social-emotional and motor development, if not also executive function and approaches to learning. Analyses of item functioning within IDELA have found that all four core domains provide related but distinct information about children's skill development (Wolf *et al.*, 2016). Also, rich evidence details the importance of non-academic skills, like social-emotional development, especially when children are

22. For more detail on the Global Fairness Initiative in Nepal, see: www.globalfairness.org/our-work/our-programs/better-brick-nepal
23. For more detail on RISE in Egypt, see: www.riseegypt.org/
24. For more detail on the Open Society Foundation, see: www.opensocietyfoundations.org
25. For more detail on the International Step by Step Association, see: www.issa.nl/
26. For more detail on the Roma Education Fund, see: www.romaeducationfund.hu/

exposed to adverse environments or life events (Betancourt *et al.*, 2014; Izard *et al.*, 2008; Heckman, 2006). Therefore, holistic assessments allow us to better understand all of the resources and needs children bring to a new learning environment, including those in the most marginalized situations. This multidimensionality has contributed to the use of IDELA within programmes supporting children who are found at the BoP from their experiences with war or disaster, like those living in refugee camps in Lebanon and Jordan, children living in earthquake-affected areas of Nepal, and orphaned or vulnerable children in rural Malawi.

Figure 8.6 Early learning and development among 5-year-old children is supported by play/learning activities at home



Source: Save the Children, 2017.

Turning to learning environments, the IDELA Home Environment Tool gathers information from caregivers about multiple dimensions of children's lives, such as poverty, home learning resources (reading materials and toys), and home learning activities (discipline and learning/play interactions). This data enables equity analyses as well as investigations of factors that influence learning. Using new data from Save the Children sites in 2017 we are investigating the relationship between adversity and early learning and development, as well as disability and development in these early years. Our analyses across sites show that the relationship of these factors to learning varies, and needs consideration context by context. That said, collecting the same data across different settings also allows for uncovering consistent trends

that have implications for programming. For example, the importance of supportive home learning environments has been documented across numerous Save the Children programme areas (*Figure 8.6*).

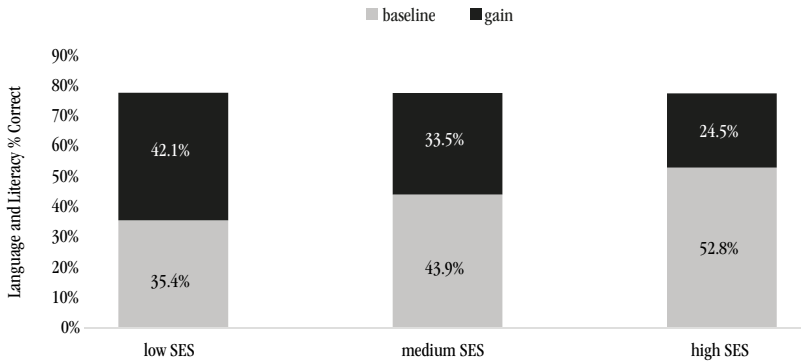
The stronger the presence of learning and play activities at home – such as reading, telling stories, singing, playing, drawing, teaching letters, numbers, or other new things – the more advanced children's learning and development. Whether at a national or subnational level, IDELA aims to capture variation in children's early learning and development, and in their early learning environments. We utilize the associated caregiver survey to develop impact and equity evidence that drive both programming and advocacy.

ECCD interventions at the bottom of the pyramid

In 2011, Save the Children critically reviewed its own ECCD practice globally, alarmed by the rise of a narrow perspective on ECCD as serving to promote emergent literacy and numeracy via rote recitations of the alphabet, choral counting, and workbooks done in seated rows. Save the Children's ECCD leadership prioritized interventions that would instead focus on holistic and age-appropriate child development. In this vision, emergent language and literacy meant storytelling, word games, print and books to promote speaking, listening, alphabet knowledge, rhyming, and knowing that print can carry meaning. Similarly, patterns and sorting, counting, shapes and problem-solving pave the way for adding, subtracting, multiplying and dividing. Importantly, self-awareness and interpersonal skills like team-building, communication, and cooperation come through opportunities to engage in these activities via play in groups, pairs, and individually. This focus on play and interaction meant to take ECCD interventions beyond academics and to support a breadth of foundational skills led Save the Children to develop the ELM (Emergent Literacy and Math) Toolkit – a set of resources for reading, math, play, and cooperative games that support emergent literacy and math skills alongside physical and social-emotional development.

ELM has a version for use in ECCD centres and preschools, as well as a version for use at home. In the first several sites, these were used together. Building on early and positive impact results from Bangladesh in 2011 (Aboud and Hossain, 2011), ELM was launched in Ethiopia and displayed its first BoP equity-promoting results in 2013 (*Figure 8.7*).

Figure 8.7 Ethiopia: average emergent language and literacy baseline and gains among children in ELM centres by SES



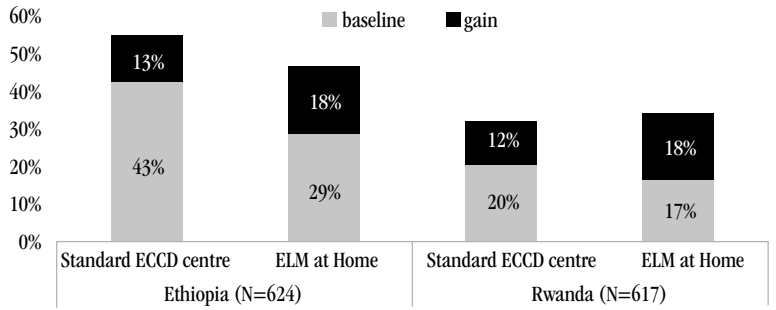
Source: Dowd *et al*., 2016.

While children in ELM centres with lower SES scores started off with lower average baseline emergent language and literacy scores than to their wealthier peers, by endline, even after controlling for parental education, age, and sex, they gained more skills in this domain than those in the higher SES quintiles. This meant that there were no differences between children's learning and development skills associated with SES as these children were about to transition into primary school.

While this finding was encouraging, the results of the larger study, which compared children receiving the ELM programme through ECCD centres with children without access to ECCD at all, gave Save the Children pause: were we really at the BoP? While we were proud to have innovated and supported children in ELM ECCD centres to score nearly 80 per cent correct on average across the foundational skill domains, children in the control group who had no access to ECCD scored an average 28 per cent correct at baseline and gained very, very little by endline five months later.

Building on these findings, the next ELM trials aimed more squarely at the BoP. The home-based version of ELM was tested alone in Ethiopia and Rwanda, where data were showing a similar need. In both sites, ELM at Home aimed to teach and empower parents to engage in play-based literacy and math activities because their children did not have access to ECCD in centres – either there was no centre near their home or they could not afford the centre fees. In both sites, children in the ELM at Home group made greater progress than peers in standard centre-based ECCD programmes (*Figure 8.8*).

Figure 8.8 With high-quality stimulation at home, children make comparable gains to those enrolled in ECCD centres



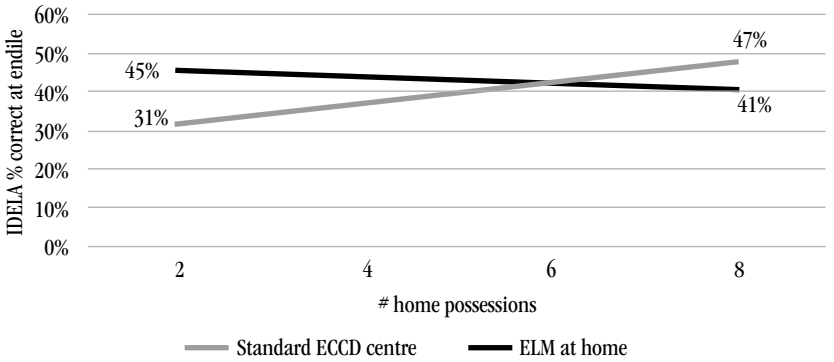
Source: Dowd *et al*., 2016.

In Rwanda, few background characteristics among the children in the ELM at Home group (i.e. age, sex, mother-tongue education, SES, and home learning environment) were found to relate significantly to learning gains. It suggests that the programme benefited children equally. This is especially important because it also suggests that caregivers of different educational and economic backgrounds were able to internalize and use the messages they received from the ELM at Home sessions. In contrast, in the standard ECCD centres children with higher SES were learning more than their peers with less family resources (*Figure 8.9*). In this model, the relationship between SES and ELM is not significant, but for standard ECD centres it is positive and significant.

The Rwanda study followed children into Grade 1 (*Figure 8.10*). It shows that children in the ELM at Home programme sustained larger longer-term gains in learning and development than their peers, and replicates the finding that those at the BoP, without any ECCD support, struggled the most as they entered primary school.

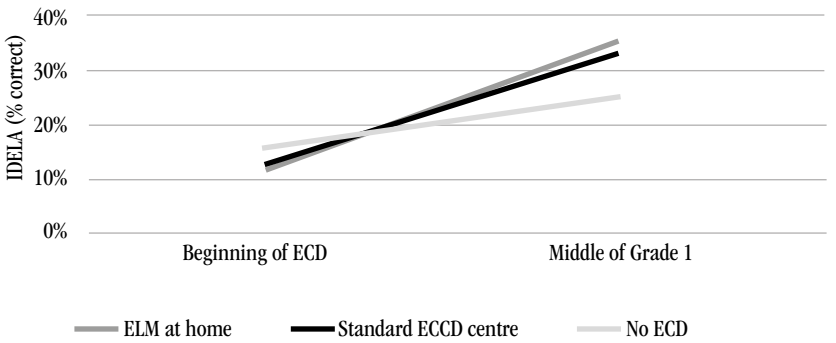
Ethiopia and Rwanda are not unique cases; many LMIC governments are aiming to provide more ECCD centre access. However, mobilizing the resources to reach all children with such services will take years. In the meantime, millions of children at the BoP could miss the opportunity to benefit from high-quality early learning experiences. We must apply good measurement tools to embrace, test, replicate, scale, and retest alternative models of ECCD, and avoid the potential to drive towards building preschools instead of building strong early learning environments for children wherever they are.

Figure 8.9 Rwanda: children benefitted equally from ELM at home, but standard ECCD centres brought greater benefit to those with high socio-economic status



Source: Save the Children, 2015.

Figure 8.10 Children receiving home-based ECCD support gain and maintain as much learning as those in ECCD centres (n=313)



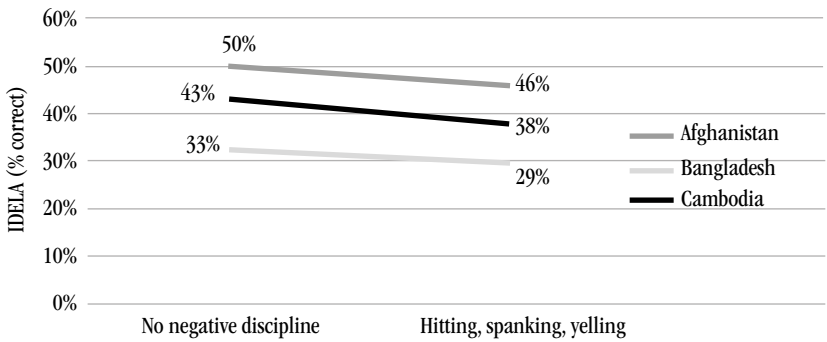
Source: Save the Children, 2015.

Challenges and limitations in ECCD at the bottom of the pyramid

There are three primary challenges in this work that relate to multidimensionality and comparability. The multidimensional nature of ECCD work presents both evidence-building and measurement challenges. From an evidence-building standpoint, we have yet to

sufficiently tackle the impact of multisectoral programming on early learning and development. For example, most livelihoods and poverty programmes assume that programming benefits children, but do not test this. Protection programming often aims to reduce family separation and promote positive parenting, but feedback on the impact of these activities on child development remains scarce and descriptive. For example, a simple view of the negative relationship between harsh discipline and children's learning and development in three sites is evident, even after controlling for parent education, SES, and the presence of positive home learning activities (*Figure 8.11*).

Figure 8.11 Negative relationship between harsh discipline and children's learning and development



Source: Save the Children, 2017.

While this is a start, we can leverage better measurement to address these gaps and hone our understanding of factors influencing early learning and development at the BoP. This is especially important to do longitudinally as the evidence about what works to sustain skill gains over time in LMICs is lacking.

On the measurement front, as noted above, approaches to learning and executive function remain less well-quantified to date. Within the ECCD sector, how to develop programmes to promote these skills is not as clear as within domains like literacy, math, motor, and social-emotional development, so incentives to develop reliable measures have been low. However, research suggests that cognitive processing abilities and learning approaches are the first line of skills affected by things like health complications and chronic stress. Thus, they may be critical for the future of cross-sectoral programming and multidimensional measurement.

Figure 8.12 Countries in which IDELA has been used to measure early learning and development in sites at the bottom of the pyramid

Source: Save the Children, 2017.

Tackling comparability, a main challenge is posed by the fact that IDELA and other tools for measuring early learning and development aim to capture a wider variety of skills and domains than most internationally comparable assessments. With this wider 'skill sample' (cf. Wagner and Castillo, 2014), IDELA produces locally (project site) if not nationally actionable evidence. This is clear from the variety of examples herein, as well as the other Save the Children and partner sites across the globe in which IDELA has been used for this type of programme and policy work (see map in *Figure 8.12*).

Promoting learning within these contexts – most being at the BoP – necessarily focuses upon local capabilities for asking and answering questions about policy and practice as they affect children's development. Whether IDELA or other tools can meet the drive to cross-national or even global comparability remains a question (Wolf *et al.*, 2017). However, what is clear already is that with a growing number of partners all using the same metric, we have greater potential to leverage change within a country or locality than ever before.

Conclusion

The attention and effort directed towards ECCD expansion expected in the next 15 years as a result of SDG 4 has the opportunity to be an

equalizing force for children at the BoP, but it also has the potential to exacerbate existing inequalities. Advocating, affecting, and generating evidence for early learning and development at the BoP requires going beyond unidimensional national averages; it demands disaggregation at subnational levels and attention to the multidimensionality of learning. It also requires promotion of a culture of using data and reflecting on facts to guide new directions. Tools like IDELA and others hold promise for illuminating the effects of the cross-sectoral programming that is being called for on the global level (Britto *et al.*, 2016). The global comparability of these and other tools remains to be established, but regardless of that outcome, it is clear that these tools can be powerful resources on national and subnational levels for those working and advocating for children at the BoP.

References

Aboud, F.E.; Hossain, K. 2011. 'The impact of preprimary school on primary school achievement in Bangladesh'. In: *Early Childhood Research Quarterly, 26*(2), 237–246.

Betancourt, T.S.; McBain, R.; Newnham, E.A.; Akinsulure-Smith, A.M.; Brennan, R.T.; Weisz, J.R.; Hansen, N.B. 2014. 'A behavioral intervention for war-affected youth in Sierra Leone: a randomized controlled trial'. In: *Journal of the American Academy of Child and Adolescent Psychiatry, 53*(12), 1288–1297.

Britto, P.R.; Lye, S.J.; Proulx, K.; Yousafzai, A.K.; Matthews, S.G.; Vaivada, T.; Perez-Escamilla, R.; Rao, N.; Ip, P.; Fernald, L.C.H.; MacMillan, H.; Hanson, M.; Wachs, T.D.; Yao, H.; Yoshikawa, H.; Cerezo, A.; Leckman, J.F.; Bhutta, Z.A. 2016. 'Advancing early childhood development: from science to scale. Nurturing care: promoting early childhood development'. In: *Lancet, 6736*(16), 1–13. Retrieved from: https://doi.org/10.1016/S0140-6736(16)31390-3.

Dowd, A.J.; Borisova, I.; Amente, A.; Yenew, A. 2016. 'Realizing capabilities in Ethiopia: maximizing early childhood investment for impact and equity'. In: *Journal of Human Development and Capabilities, 17*(4), 477–493, Retrieved from: DOI: 10.1080/19452829.2016.1225702.

Hart, B.; Risley, T.R. 1995. *Meaningful differences in the everyday experience of young American children*. Baltimore, Md.: Paul H. Brookes.

Harvard University Center for the Developing Child. 2016. *From best practices to breakthrough impacts: a science-based approach to building a more promising future for young children and families*. Retrieved from: https://bit.ly/29gtnwe

Heckman, J.J. 2006. 'Skill formation and the economics of investing in disadvantaged children'. In: *Science, 312*(5782), 1900–1902.

Heckman, J.J.; Masterov, D. 2004. 'Productivity argument for investing in young children', Working Paper 5. Chicago, Ill.: Invest in Kids Working Group Committee for Economic Development. Retrieved from: http://jenni.uchicago.edu/Invest/FILES/dugger_2004-12-02_dvm.pdf

Izard, C.E.; King, K.A.; Trentacosta, C.J.; Morgan, J.K.; Laurenceau, J-P.; Krauthamer-Ewing, E.S.; Finlon, K.J. 2008. 'Accelerating the development of emotion competence in Head Start children: effects on adaptive and maladaptive behavior'. In: *Development and Psychopathology, 20*, 369–397.

McCoy, D.C.; Peet, E.D.; Ezzati, M.; Danaei, G.; Black, M.M.; Sudfeld, C.R.; Fawzi, W.; Fink, G. 2016. 'Early childhood developmental status in low- and middle-income countries: national, regional, and global prevalence estimates using predictive modeling'. In: *PLoS Medicine, 13*(6): e1002034.

Nores, M.; Barnett, W.S. 2010. 'Benefits of early childhood interventions across the world: (under) investing in the very young'. In: *Economics of Education Review, 29*, 271–282.

Oxford Poverty and Human Development Initiative. 2016. Global MPI Interactive Databank. Retrieved from: www.dataforall.org/dashboard/ophi/index.php/

Patrinos, H. 2017. 'Investing in the early years and early reading'. *Huffington Post,* 15 February. Retrieved from: www.huffingtonpost.com/ entry/ investing-in-the-early-years-and-early-reading_us_589fc1e0e4 b0e17 2783a9d4b?ncid=engmodushpmg00000004

Pisani, L.; Borisova, I.; Dowd, A.J. 2015. 'International development and early learning assessment', technical working paper. Washington DC: Save the Children. Retrieved from: http://resourcecentre.savethechildren.se/library/international-development-and-early-learning-assessment-technical-paper

Save the Children. 2015. *Early Literacy and Maths Initiative (ELMI) Rwanda endline report*. London: Save the Children. Retrieved from: http://resourcecentre.savethechildren.se/sites/default/files/documents/rwanda_elmi_endline_report_final.pdf

——. 2016. *National ECD impact evaluation study*. Bhutan: Save the Children. Retrieved from: https://resourcecentre.savethechildren.net/sites/default/files/documents/bhutan_nationa_study_endline_final.pdf

——. 2017. *Windows into early learning and development: cross country IDELA findings fueling progress on ECD access, quality and equity*. London: Save the Children. Retrieved from: https://i.stci.uk/sites/default/files/libraries/IDELA%20REPORT.pdf.

Wagner, D.A.; Castillo, N.M. 2014. 'Learning at the bottom of the pyramid: constraints, comparability and policy in developing countries'. In: *Prospects, 44:* 627. Retrieved from: DOI 10.1007/s11125-014-9328-8.

Walker, S.P.; Chang, S.M.; Powell, C.A.; Grantham-McGregor, S.M. 2005. 'Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study'. In: *Lancet, 366*(9499), 1804–7. Retrieved from: https://doi.org/10.1016/S0140-6736(05)67574-5

Wolf, S.; Halpin, P.; Yoshikawa, H.; Pisani, L.; Dowd, A.J. 2016. *Assessing the construct validity of Save the Children's International Development and Early Learning Assessment (IDELA)*. Washington DC: Save the Children. Retrieved from: http://resourcecentre.savethechildren.se/library/assessing-construct-validity-save-childrens-international-development-and-early-learning.

Wolf, S.; Halpin, P.; Yoshikawa, H.; Pisani, L.; Dowd, A.J.; Borisova, I. 2017. 'Measuring school readiness globally: assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDELA) in Ethiopia'. In: *Early Childhood Research Quarterly, 40*(4), 21–36.

Commentary

Abbie Raikes

Children at the bottom of the pyramid face unique and powerful barriers to achieving their potential (Black *et al.*, 2017). Below, three areas for research are outlined.

First, while much can be learned from cross-sectional studies, more longitudinal studies are needed to document how early childhood spent at the BoP affects lifelong learning. Learning begets learning – meaning that children with stronger skills also learn more over time, leading to increased competence for some children, and increasing gaps in learning for other children, thus cementing inequity. The question becomes, what helps close this gap?

Evidence on the link between early childhood skills and later learning is strong in high-income countries, but the role of specific skills and competencies in influencing later achievement has not yet been explored in many low-income countries. For example, self-regulation facilitates faster growth in math and literacy skills throughout the primary school years in samples in the USA (McClelland, Acock, and Morrison, 2006), and early math may have a strong role in encouraging later reading (Duncan *et al.*, 2007). More detailed study of multidimensional learning among children at the BoP may contain clues for closing the gap. Likewise, home environments have powerful and lasting effects on children's learning (e.g. Melhuish *et al.*, 2008). But in many countries, the potentially powerful impact of supporting parents to provide stimulating home environments is not yet fully leveraged. Deeper understanding of patterns and influences on child development is essential for effective interventions.

Second, poor health and nutrition, and exposure to violence in early childhood, are rarely mentioned when discussing later learning. Malnutrition, for example, while emerging as an influence on children's development between ages of birth and 2 years, affects learning throughout the life span (Alderman, Hoddinott, and Kinsey, 2006). Results from interventions focused on very young children indicate that stimulation can help ameliorate the effects of undernutrition on cognition (e.g. Grantham-McGregor *et al.*, 1991), but few educational interventions in the preschool or early school years specifically address the effects of early undernutrition on learning. Exposure to violence and poor health, including infectious diseases, have profound implications

for neurological development, beginning in early childhood and extending throughout schooling. Addressing health, nutrition, and exposure to violence in early childhood and beyond may support learning at the BoP.

Finally, a comprehensive research strategy is needed to develop reliable and cross-culturally relevant metrics of early childhood development, especially if there is desire to compare across countries. It is essential to capture a broad range of skills (not only for early childhood, as a wide range of skills has been shown to influence later learning too). While it is unclear whether a singular, cross-culturally valid metric of early childhood is possible, systematic data on early childhood development is important for tracking inequity. A coordinated and cooperative research agenda should be developed to inform the development of metrics, including longitudinal data on multiple aspects of young children's development, attention to the pros and cons of generating comparable data, and the acknowledgement of many factors that affect learning, including family environments, health, nutrition, and exposure to violence.

References

Alderman, H.; Hoddinott, J.; Kinsey, B.; 2006. 'Long term consequences of early childhood malnutrition'. In: *Oxford Economic Papers*, *58*(3), 450–474.

Black, M.M.; Walker, S.P.; Fernald, L.C.; Andersen, C.T.; DiGirolamo, A.M.; Lu, C.; McCoy, D.C; Glink, G; Shawar; Devercelli, A.E. 2017. 'Early childhood development coming of age: science through the life course'. In: *Lancet*, *389*(10064), 77–90.

Duncan, G.J.; Dowsett, C.J.; Claessens, A.; Magnuson, K.; Huston, A.C.; Klebanov, P.; Pagani, L.S.; Feinstein, L.; Engel, M.; Brooks-Gunn, J.; Sexton, H.; Duckworth, K.; Japel, C. 2007. 'School readiness and later achievement'. In: *Developmental Psychology*, *43*(6), 1428.

Grantham-McGregor, S.M.; Powell, C.A.; Walker, S.P.; Himes, J.H. 1991. 'Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study'. In: *Lancet, 338*(8758), 1–5.

McClelland, M.M.; Acock, A.C.; Morrison, F.J. 2006. 'The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school'. In: *Early Childhood Research Quarterly*, *21*(4), 471–490.

Melhuish, E.C.; Phan, M.B.; Sylva, K.; Sammons, P.; Siraj-Blatchford, I.; Taggart, B. 2008. 'Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school'. In: *Journal of Social Issues*, *64*(1), 95–114.

Commentary

Kathy Hirsh-Pasek and *Andres S. Bustamante*

We agree with Dowd and Pisani who state that 'early learning and development at the bottom of the pyramid requires going beyond unidimensional national averages'. If we are to educate young children in ways that enable them to learn skills for a 21st-century workforce, we must abandon myopic approaches that focus on a small set of competencies culled into a 'representative' score. This preoccupation with sole outcome scores in reading and math has led to interventions that at best nudge rather than shove the proverbial needle forward. It is time for a new approach that focuses on a breadth of skills, a breadth of measurements, and a breadth of contexts that enable children at the BoP to reach their fullest potential.

The call for breadth of skills is evident in both the scientific and the policy literatures. Reviewing thousands of articles, Golinkoff and Hirsh-Pasek (2016) argue for a dynamic suite of skills that include collaboration, communication, problem-solving, creativity as well as content, and confidence (grit), that together, not singly, predict trajectories for children in school and in life. It is well established that while math and language skills at kindergarten entry are predictive of children's long-term achievement (e.g. Yoshikawa *et al.*, 2013), so too are executive function skills (e.g. McClelland *et al.*, 2013), approaches to learning (e.g. Bustamante *et al.*, 2017*a*), and social-emotional skills (e.g. Jones *et al.*, 2015). These skills are evidence-based, malleable, and measurable, making them perfect targets for intervention.

Breadth of measurement poses an interesting challenge as we endorse a broader definition of learning. How can we measure a profile of skills in a cost-effective and meaningful way? Our measurement tools are inadequate, even the well-worn tests with well-understood constructs. The PPVT language test, for example, predominantly assesses noun learning rather offering a portrait of vocabulary that investigates nouns and verbs and the spatial language. Further, focusing only on vocabulary growth is itself inadequate for explaining reading and math outcomes. Language-learning – the crucible for reading comprehension – is broader than mere word learning. As Dowd and Pisani note, finding good measurement is a challenge for the field, and one that we must embrace for those areas in which we think we have strong accountability (e.g. reading).

Breadth of context offers the third leg of the argument for a more comprehensive approach to intervention. As Dowd and Pisani demonstrate, home interventions can be as effective, if not more so, than school interventions. In the USA, children spend only 20 per cent of their waking time in school (Meltzoff *et al.*, 2009) and we must address how they spend the other 80 per cent. Research shows convincingly that high-quality interventions that begin early (age 0–3) and engage parents show consistent, long-lasting effects on child outcomes (Bustamante *et al.*, 2017*b*).

Scientific evidence offers a clear mandate that we expand our approaches as we help all children learn. Even brain data shows that while poverty has real and early effects on brain development (Hair *et al.*, 2015; Lawson *et al.*, 2013), it does not alter the way in which children learn. All children – at the bottom and top of the pyramid – learn best when they are active and engaged, through a breadth of skills and in multiple contexts in and out of school. Those who suggest that we must focus on a set of narrowly prescribed skills taught in an atmosphere of harsh discipline are not applying mounting evidence from the newly amalgamated field of the science of learning. It is time to change our approaches so that – to borrow from Shonkoff and Fisher (2013) – we adopt a broader approach to educating all children while asking, 'What works for whom and why, and in what contexts?' Only then can we hope to really move the needle far enough to make lasting differences in child outcomes.

References

Bustamante, A.S.; White, L.J.; Greenfield, D.B. 2017. 'Approaches to learning and school readiness in Head Start: applications to preschool science'. In: *Learning and Individual Differences*, *56*, 112–118.

Bustamante, A.S.; Hirsh-Pasek, K.; Vandell, D.H.; Golinkoff, R.M. 2017. 'Why DeVos should embrace early childhood education'. In: Brookings Institute Education blog. Retrieved from: www.brookings.edu/blog/education-plus-development/2017/02/09/why-devos-should-embrace-early-childhood-education/.

Golinkoff, R. M.; Hirsh-Pasek, K. 2016. *Becoming brilliant.* Washington DC: APA LifeTools.

Hair, N.L.; Hanson, J.L.; Wolfe, B.L.; Pollak, S.D. 2015. 'Association of child poverty, brain development, and academic achievement'. In: *JAMA Pediatrics*, *169*(9), 822–829.

Jones, D.E.; Greenberg, M.; Crowley, M. 2015. 'Early social-emotional functioning and public health: the relationship between kindergarten social competence and future wellness'. In: *American Journal of Public Health, 105*(11), 2283–2290.

Lawson, G.M.; Duda, J.T.; Avants, B.B.; Wu, J.; Farah, M.J. 2013. 'Associations between children's socio-economic status and prefrontal cortical thickness'. In: *Developmental Science, 16*(5), 641–652.

McClelland, M.M.; Acock, A.C.; Piccinin, A.; Rhea, S.A.; Stallings, M.C. 2013. 'Relations between preschool attention span-persistence and age 25 educational outcomes'. In: *Early Childhood Research Quarterly, 28*(2), 314–324.

Meltzoff, A.N.; Kuhl, P.K.; Movellan, J.; Sejnowski, T.J. 2009. 'Foundations for a new science of learning'. In: *Science, 325*(5938), 284–288.

Shonkoff, J.P.; Fisher, P.A. 2013. 'Rethinking evidence-based practice and two-generation programs to create the future of early childhood policy'. In: *Development and Psychopathology, 25*, 1635–1653. Retrieved from: DOI 10.1017/S0954579413000813.

Yoshikawa, H.; Weiland, C.; Brooks-Gunn, J.; Burchinal, M.R.; Espinosa, L.M.; Gormley, W.T.; Zaslow, M.J. 2013. 'Investing in our future: the evidence base on preschool education'. Ann Arbor, Mich.: Society for Research in Child Development.

Chapter 9
Learning progression models that support learning at the bottom of the pyramid

Esther Care, Pamela Robertson, and *Marlene Ferido*

One of the characteristics associated with normative assessment approaches is that they pay allegiance to the notion that there is a natural order to society, with some members located in the upper reaches and others in the lower. This notion influences how assessments are designed and used such that they reflect the normative expectation that a small number of students will achieve at peak level, a large group of students will be average, and a predetermined number will fail. This model has been shown to reflect many human characteristics reasonably well, and has been used as a tool in selection situations where the nature of the environment only allows for a limited number of individuals to progress.

Alignment of assessment approach with 21st-century aspirations

This model is not, however, aligned with the expectations of the Sustainable Development Goals, which call for *all* children to have access to educational opportunities and to be able to benefit from them (access *plus* learning). This aspiration challenges those operating assumptions that require individuals to be located across a normal distribution in terms of learning outcomes. In some contexts, and on some occasions, a pass/fail approach to assessment might be a reasonable strategy. However, where the purpose of assessment is to inform and improve learning outcomes, the pass/fail approach can lead to further marginalization of at-risk children and youth. We need assessment approaches that inform and guide students' learning progress. In order to optimize the potential of assessment, perceptions about the use of assessment need to change, and the assessment tools themselves need to be designed in a way that generates data usable for and relevant to all students.

In response to SDG 4, education systems need to review their curricular, teaching, and assessment tools to ensure these cater to all students. Taking the perspective that learning consists of acquiring increasingly complex sets of skills and knowledge means that it is viable to describe each and every student's current level of learning in such a

way as to identify what they are ready to learn next. Assessment tools based on this perspective can capture the capabilities of all students in a way that can support their learning, as opposed to assessment tools which function primarily as a divider of the learner population into those who can and those who can't.

The majority of us accept the concept that we need to master simple skills before moving on to more complex ones. But some do not. One set of skills that exemplifies some controversy around this is literacy. Over the past few decades we have seen a shift from a phonetics approach to teaching of basic literacy to whole word, and not quite back again, but to a middle ground. What we have learned from that particular example is that we need to accompany deconstruction and simplification of skills with meaning. Notwithstanding such skirmishes, our education systems are based absolutely on the notion that one proceeds from the more basic to the more advanced, and that mastery of the basic is a precondition for progress. It makes sense therefore that assessment should reflect this educational perspective. And more than that, in order to know at what point to intercede with a student in terms of their learning, we need to evaluate both what the student knows and what they are 'ready' to learn. The answer to this need is assessment.

Assessment that is constructive is not about providing or denying students access to further education or opportunities. It is not about pass or fail. Well-designed assessment has the capacity to provide information that has a great deal more utility. It can identify what appears to be well known or understood as well as what a student is currently on the cusp of learning, and that information can be used to guide instruction. Assessment also has the potential to inform an education system about how it is performing relative to its goals for its learners. This summative use of assessment data can be a primary source for the monitoring and evaluation of the system. Not only can it be used to identify where particular strengths and weaknesses of overall performance exist across regions or specific population groups, it can also be used to analyse specific within-discipline strengths and weaknesses across and within the system – with consequent opportunity for targeting of areas for improvement in curricula, curricular resources, or teacher development. The assessment data reported in this article is used for this latter purpose, and can equally be used to inform individual teaching interventions. It should be recognized that not all assessments have the capacity to inform both functions. However, assessments that are developed from a coherent blueprint that structures the learning domain of interest, with

items that clearly indicate the content specified by the learning domain, and that are appropriately scored, can provide such opportunity.

Learning readiness

The setting of assessment tasks for students requires knowledge of the domain that is being assessed (e.g. literacy), knowledge of how performance on that domain can vary from basic competency through to more sophisticated mastery, and estimates of students' 'zone of proximal development' (Vygotsky, 1978) or general functioning level so that the assessment can be appropriately targeted. The zone of proximal development is the point at which the student is most ready to learn, and at which intervention will have the greatest impact; it is the area between the tasks a student can do without assistance and those where the student requires help. When a large percentage of students cannot pass a test, it is an indication that the test is not well targeted – it does not present items or questions with which many students can engage because the items do not fall within the student's zone of proximal development. In other words, it does not have the appropriate range of difficulty that is matched to current student competence. What this means is that the test result tells one very little about what the student is capable of, has learned, or is ready to learn. This scenario obtains regardless of whether we are talking about a classroom-based assessment task or an assessment event that is scheduled across many students. This approach is criterion-referenced where student responses to test items help to identify directly what is known or achievable, rather than placing a student at a rank within a population.

When a test that is in principle constructed to assess a particular target population in fact is not able to provide a similar amount of meaningful information about a particular group of students within that population, that test can be seen as biased. Often the concept of bias is discussed in the context of gender, race, and language. However, bias is equally an issue when a particular group within an overall population distribution is prevented from being able to provide useful data, perhaps because they are below the floor or above the ceiling of the test. From an ethical angle, assessments must be inclusive of all distribution-locatable groups within the target population. If the target population is all students at a particular age/grade level in formal education, similar levels of information about all of these students should be accessible through the assessment, regardless of where they are located in the overall distribution of all students on the characteristic or capacity of interest.

Many of our assessment approaches reflect a time when there was limited opportunity for students to progress all the way through a formal learning system. Assessment was one device used to support this reality, such that only students who showed academic promise were promoted through the system. Normal distribution assumptions provided a useful structure, both philosophically and technically, to support this situation. However, where the intention of an education system is to maximize the number of students who progress all the way through basic and secondary education, both assessment and pedagogical strategies need to support this 21st-century perspective.

Assessment for all students

The 'base of the pyramid' is where all students start in terms of their learning. The image of the BoP is that it is close to the ground, stable and solid, providing the foundation from which to build. This foundation continues to provide stability; it is returned to in order to consolidate, review, and then progress. Learning at the base informs all later learning, and is inherent in the knowledge and skills we acquire over time. In the assessment context, a test item incorporates the prerequisite knowledge and skills acquisition trail in its targeting the construct of interest.

In the context of the SDG focus on learning for all, we should no longer be interested in whether a student passes or fails. Rather, we should be interested in facilitating their learning. This means we need assessment tools that provide this information rather than dividing any given group into pass and fail. Concomitantly, as curricula become more complex, the knowledge sequence in the learning area also must change. This should alert us to relevant access points for the development of linked assessments. Such an assessment can better pinpoint where the learner is located on a learning progression from the base upward. However, in order for this to occur, test items must be written to target the range on the learning progression to the likely range of capability of the learners for whom they are intended.

Developing assessment tools that sample both prerequisite skills and learning goals within particular subject areas provides the opportunity for identification of the functioning levels of all students within a grade level. If test items that are similarly difficult can be analysed and found to be targeting particular types and levels of skills, the descriptions of these can then be matched to the student learning abilities. Using this approach, student results can be reported by text that describes the learning achievement, and both students and teachers are

able to use the results. Identification of groups of students at particular competency levels enables the teacher to address their learning needs more efficiently (Care *et al.*, 2014).

Example of approach

A recent example of this approach to assessment was developed by the Assessment Curriculum and Technology Research Centre (ACTRC; www.actrc.org) for use by the Philippines Department of Education. In the K–12 (from kindergarten to Grade 12) education reform programme of 2013, among other initiatives, a new curriculum for science was implemented in schools across the country. Through this curriculum, the aim is to develop scientific literacy among students so that they are able to make judgements and decisions on the applications of scientific knowledge that may have significant impact in everyday life (Care *et al.*, 2014). The new curriculum is designed according to three domains of learning science: (1) understanding and applying scientific knowledge, (2) performing scientific processes and skills, and (3) developing and demonstrating scientific attitudes and values. The Philippine curriculum includes statements outlining the progression of science inquiry skills and expectations of the rate at which students will develop these skills, addressing some of the recommendations resulting from a comparison of the Philippine curriculum with other countries of interest (Care and Griffin, 2011). Scientific concepts and skills are now presented by the education system through increasing levels of complexity from one grade level to another in a spiral progression that provides opportunity for development of deep understanding of core concepts.

ACTRC's assessment-based research study was designed to investigate the progress of students' conceptual knowledge and skills for each unit of chemistry (here the study of 'Matter') over the four years of the junior secondary curriculum. Identification of progress was achieved by assessing students in Grades 7–10. Determining what students know and can do at various stages can be used both to inform subsequent teaching, and to assist with future reviews of the science curriculum.

An important feature of this research is that students are tested directly before they enter each unit of chemistry. The rationale for this approach draws attention to depth of student learning, as opposed to knowledge acquisition or surface learning. The approach provides an indication of the skill level of students prior to each relevant quarter (or school term), and the level of skill retained from the previous relevant quarter (school term). The skill level is indicative of the knowledge,

skills, and understandings retained long-term from previous units of the subject. Of interest is what the students have acquired and retained, from those who demonstrate very little, to those whose knowledge and skills appear to be well aligned with curriculum expectations. Across the full life of the study, five linked tests were developed in order to locate students across the grade levels of interest on the same scale. Some common items link the tests of approximately 50 items each. The use of common items enables the comparison of student acquisition of skills and knowledge across different grade levels. The process of test development includes initial curriculum mapping and development of test blueprint, item development, item review, pilot, analysis of pilot data, and then selection of items for the final test form.

The blueprint to structure the development of test items was created using information from a curriculum audit. This included what concepts and skills students need to have in order to access a specific grade-level chemistry curriculum. The blueprint took into account the knowledge strands that run through the different grades and those that appear only at one or some of the grades, as well as the relative importance of the categories/strands at each grade level.

To articulate the skills integral to each grade level of the curriculum, specific behaviours that a student might be expected to demonstrate and that could be demonstrated in a pen and paper test were identified. Where the same behaviours appear at multiple grade levels, these were noted at each relevant level. The behaviours were also classified as 'essential' or 'advantageous' by curriculum experts in the context of their achievement to inform the next grade level.

For the main study, the test data collected from the student sample was calibrated using the one-parameter simple logistic model (Rasch, 1960) and student ability estimates were produced. The benefit of Rasch modelling is that the students and items can be placed on the same scale. This enables the comparison of students and the concepts and skills assessed by the test. The relative positions of items and students are dictated by the data fit to the Rasch model. When an item is positioned 'next to' a student on a visual map of the data, the student has a 50 per cent chance of answering that item correctly. When the item number is linked with the concept or skill assessed by the item, the horizontal visual alignment identifies which concepts or skills the aligned students are ready to learn. For usability, clusters of similar skills are identified, and a level description is written to encapsulate the main ideas that students aligned with each cluster are ready to learn.

Pre-Grade 7

For Grade 7, the test blueprint includes the prerequisite concepts and skills considered necessary for students to engage in the Grade 7 Matter curriculum. These prerequisites were searched for in the Grade 3–6 curricula. It should be noted that some of these skills are not explicitly stated within these curricula. Despite this, the prerequisites were included in the test blueprint due to the Grade 7 curriculum demand.

The pre-Grade 7 test conceptually and empirically separates into six levels (A through F). These levels are described in *Figure 9.1*. Levels A to D contain the prerequisite concepts and skills for Grade 7 Matter. These include concepts such as identifying examples and properties of solids, liquids, and gases, and recognizing changes in state such as melting and freezing, as well as skills such as drawing conclusions from results and recognizing scientific aims. Since students in Levels A to D are identified as 'ready to learn' these concepts (e.g. that a gas cannot be held in one's hand), it means that they have not yet mastered the prerequisites for the Grade 7 Matter curriculum. Given that these students are located across Levels A to D, obviously some have mastered more of the assessed skills than others. The concepts and skills in Levels E and F are beyond those identified by the curriculum experts as being prerequisite knowledge. For example, linking different states of matter to the particulate nature of the substance is advantageous, but not a prerequisite for Grade 7 Matter. Similarly, the skill of controlling variables is taught as part of scientific method within Grade 7, so students at Level E are ready to learn this Grade 7 material. This implies that students in both Levels E and F are adequately prepared to engage in the Grade 7 Matter curriculum.

Figure 9.1 is a person–item map, sometimes referred to as a variable or Wright Map (Wilson, 2004). It shows two vertical histograms, providing a schematic representation of how item difficulties and student abilities relate to each other with respect to an underlying construct. The left side shows students as a series of Xs, with each X representing 2.8 students, while the right side shows items, labelled by number. The distribution of the students is from most able at the top to least able at the bottom. The items on the right side are distributed from most difficult at the top to least difficult at the bottom. Theoretically, when students and items are opposite each other on the map, the difficulty of the item and the ability of the student are comparable, so the student has a 50 per cent probability of answering the item correctly. Therefore items that appear very low

down on the distribution have a much higher probability of being responded to correctly by a much higher proportion of the students, and items high on the distribution have increasingly low probabilities of being responded to correctly. Accordingly, this item-response-theory-derived map makes it possible to order students and items on the same scale. (For a full discussion of the use of Wright Maps in the context of educational assessment, see Black, Wilson, and Lao, 2011).

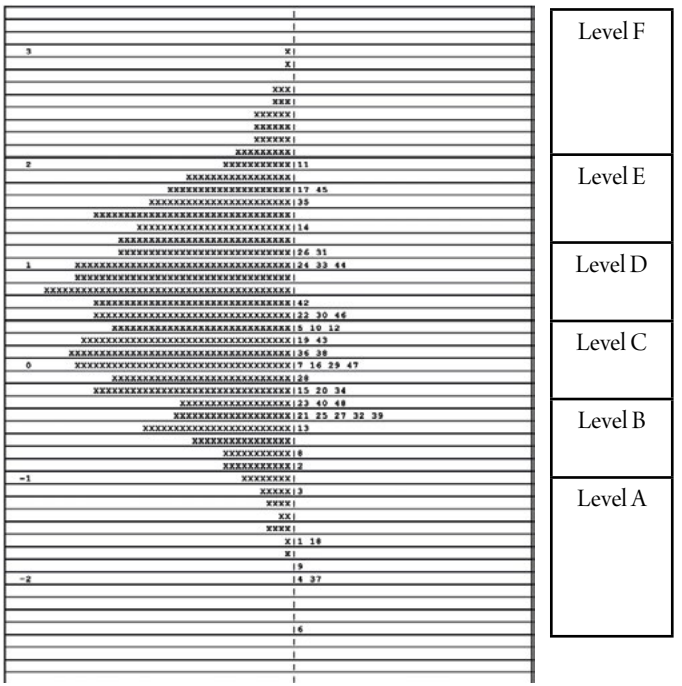
Analysis of the content and skills of the items provides experts with information about clusters of skills, each of which are qualitatively more difficult than the previous, and using this information, broad levels of skills from Levels A to F were identified. As can be seen, the majority of students are located across Levels A to D. This finding raises an obvious issue; the students are not ready to learn at the Grade 7 level. Concepts, content, and skills that target the Grade 7 level will not target these students' zone of proximal development. If students do not receive appropriate intervention in their zones of proximal development, the provision of instruction according to the grade-level curriculum is a waste of time. Not only is it not instructive, it can be assumed to influence student self-esteem, increase frustration, and decrease engagement. However, although we might assume that provision of the grade-level curriculum and instruction will not be able to be acted upon adaptively by students – hypothesizing that skills build on skills, and knowledge upon knowledge – that assumption should be questioned. The next section examines this issue.

Continuing neglect of students at the 'bottom of the pyramid'

When students at the BoP are not learning successfully at a particular grade level, the presumption by some teachers and parents is that students merely need to apply themselves or work harder in order to increase their achievement levels. In fact, no matter how hard these students work at their current grade-level curriculum, they will not be able to overcome the deficit of not having mastered earlier learning. What evidence can we provide to support this hypothesis?

If a certain proportion of students have not mastered prerequisite skills for studying at a particular year level, then it is logical that at least that proportion will not be well prepared for subsequent year levels. This assumes that a curriculum is indeed designed hierarchically, and that each year's curriculum builds on the previous one. Accordingly, data are presented on student readiness for Grades 8, 9, and 10 to determine whether this contention is reasonable.

Figure 9.1 Skill level descriptions and variable map of student and item distribution



Source: Authors' calculations; data from ACTRC Science Curriculum study, https://actrc.org. See: https://actrc.files.wordpress.com/2013/07/science-curriculum-pre-grade-7-report.pdf

Note: Alignment of skill level descriptions against the map is approximate. Level A, for example describes students from a logit of -1 downward; Level B describes students from a logit of -1 up to almost zero, and so on.
Level F: Students are beginning to understand the use of words describing the properties of matter, including mass, shape, and volume. They are learning to describe physical changes in everyday materials, e.g. that metal softens when it is heated. They are starting to choose the correct scale for accurate measurement, and demonstrating a relationship between variables graphically.
Level E: Students are learning that substances that exist in different states (ice and water) differ in terms of molecular arrangement. They are learning to identify an acid from a list of everyday substances. They are starting to identify the procedure to follow in investigating a particular property of a substance, learning how to control variables, and beginning to recognize a scientific experiment.
Level D: Students are starting to use understanding of how matter behaves to make real-life predictions; understand a gas cannot be held in the hand, some substances dissolve and others settle, a solid is not always flat. They are learning how to interpret and compare different scales, which equipment is needed for a particular experiment, how to generate explanations from results, and how to tabulate data.
Level C: Students are deepening their understanding of the properties of matter; that magnets attract iron; that electricity is a movement of electrons; that sand will not dissolve in water. They are learning to explain scientific observations, to interpret graphed and tabulated data and to identify the aim of an investigation.
Level B: Students are learning the basic properties of solids, liquids, and gases. They are beginning to know the properties and behaviour of everyday materials; that metals rust and conduct heat and electricity; that substances dissolve faster in hot water. They are starting to be able to identify elements that are not metals. They are learning the difference between an observation and an explanation and how to read tabulated data.
Level A: Students are starting to identify common solids, liquids, and gases based on definition. They may identify a metal from a list of non-metals and may have a basic understanding of some properties of everyday materials, e.g. that copper conducts electricity.

Pre-Grade 8, 9, and 10 tests were designed, piloted, trialled, and used according to the approach used for Grade 7. The percentages of students entering these grades who were 'ready to learn' prerequisite concepts and skills are shown in *Table 9.1*, including those at pre-Grade 7 who were discussed in the previous section.[27] The identification of ready to learn followed the same processes and conventions as for pre-Grade 7 but naturally the tests targeted and described successively more sophisticated concepts and skills than those depicted in *Figure 9.1*.

As can be seen, the proportion of students not ready to engage with the grade level curriculum increases each year. There is of course some attrition, with students leaving the formal education system each year, so it may well be that *Table 9.1* percentages under-represent what the situation would be if all students remained in school through to Grade 10. Notwithstanding, it is clear that there is no recovery from the problem that emerged at pre-Grade 7. Increasing proportions of students enter grade levels with little chance of being able to engage with the curriculum due to their lack of accumulation of prerequisite concepts and skills.

Table 9.1 Percentages of students ready and not ready to engage with grade-level curricula

	Ready	Not ready
Pre-Grade 7	619 (39%)	978 (61%)
Pre-Grade 8	398 (27%)*	1,085 (73%)
Pre-Grade 9	207 (13%)*	1,395 (87%)
Pre-Grade 10	153 (11%)	1,212 (89%)

Source: Authors' calculations, data from ACTRC Science Curriculum study, https://actrc.org
Note: * Significant ($p < 0.00001$) decline in readiness from previous grade.

Conclusion

The point of all this is that the tests in this study are designed to identify the achievement of all students. The approach does not adopt a pass/fail perspective, it is indifferent to distributions of students according to a normative paradigm, and it provides information that benefits all students and their teachers. Most importantly, the approach provides information for system-level analysis. It identifies the specific concepts

27. Note that the data in for pre-Grade 7 include students from both regular and science-oriented high schools, whereas data for only regular high schools are included for all grade levels in *Table 9.1*.

and skills that students are not integrating, and provides a clear message to the system that the student cohorts are not able to recover from early lack of acquisition of prerequisites to continue study successfully.

Approaching assessment from this perspective clearly provides tools that serve multiple audiences – system, teacher, and student. It provides insights for immediate instruction, but also highlights at system level the obstacles faced by a teacher in a classroom who is under pressure to move through a scripted curriculum. The approach supports evaluation of every student's achievement, and does not discriminate against a group at any point of the achieved distribution, or in more educational terms, against students who might be grouped in different percentile ranges. The notion of the BoP is less impactful when assessment is designed to capture the achievements of all students – this is no more complex than ensuring that tests are designed for their targets.

The use of a normative paradigm supported the pyramid notion – that very few would achieve great things – and, we have argued, is a self-fulfilling outcome. New evidence suggests a greater opportunity for larger proportions of children to benefit from education, and more need for that to occur. We need assessment assumptions that can support a larger proportion of the population that can learn more advanced skills than was believed possible in the past. Part of that capacity to learn is optimized through knowing how to teach, and that knowing how to teach is informed by the capacity of assessment tools to reflect the realities of each student's progress. Forget the BoP – it's just the first building block. But without it being carefully constructed, the blocks above will not hold together.

References

Black, P.; Wilson, M.; Yao, S. 2011. 'Road maps for learning: a guide to the navigation of learning progressions'. In: *Measurement: Interdisciplinary research and perspectives, 9*(2–3), 71–123.

Care, E.; Griffin, P.; Zhang, Z.; Hutchinson, D. 2014. 'Large-scale testing and its contribution to learning'. In: C. Wyatt-Smith, V. Klenowski, and P. Colbert (Eds), *Designing assessment for quality learning* (pp. 55–71). Dordrecht, Netherlands: Springer.

Care, E.; Griffin, P. 2011. *Curriculum comparison study for the Philippines basic education sector.* Melbourne, Vic.: Assessment Research Centre, University of Melbourne. Retrieved from: http://trove.nla.gov.au/work/155913772?selectedversion=NBD47651089

Rasch, G. 1960. *Probabilistic models for some intelligence and achievement tests.* Copenhagen: Danish Institute for Educational Research.

Vygotsky, L. 1978. 'Interaction between learning and development'. In: L. Vygotsky (Ed.), *Mind in society* (pp. 79–97). Cambridge, Mass.: Harvard University Press.

Wilson, M. 2004. 'On choosing a model for measuring'. In: E. V. Smith and R. M. Smith (Eds.), *Introduction to Rasch Measurement.* Maple Grove, MN: JAM Press.

Commentary

Rachel Hinton

Care fills a gap in a debate that has been dominated by the drive for global reporting. She presents an assessment that facilitates learning for the *teacher* in order to focus on SDG 4 – all children learning. She warns against the 'pass' or 'fail' mentality, arguing for assessments that identify children's knowledge gaps. This enables 'teaching at the right level' – known to significantly raise learning outcomes (Banerjee *et al.*, 2016). Global and national assessment data are crucial to track the impact of reforms and equity – yet often the institutional capacity for analysis is limited and data fail to be acted upon. The Philippines Assessment Curriculum and Technology Research Centre (ACTRC) tool is a welcome outlier. It has led to action from data, and a shift of policies in favour of those learners who are at the BoP.

While Care's paper demonstrates the potential of putting data in the hands of teachers, an important issue that could be further explored is the potential impact of putting data in the hands of *parents*. Parents place great hope in educational establishments to break the poverty trap, yet many harbour false beliefs about their children's academic achievement. The Participatory Poverty and Vulnerability Assessment in Ghana reported that, without any access to information, caregivers significantly overestimated the quality of schooling (UK DFID, 2011). A study in rural Pakistan found that the provision of information on learning to parents in rural villages led to several positive effects, including improved parental knowledge, and significantly improved subsequent student test scores (Andrabi, Das, Khwaja, 2016). The ASER household assessments provide an alternative model whereby parents witness the survey being conducted, and visual analysis is shared at the local level (Banerji, 2016). Yet the impact of providing assessment scores to parents in resource-poor contexts remains under-researched, despite a 10-fold increase in systematic evaluations to measure learning (Evans and Popova, 2016).

What about *children* having access to the data? Care notes that 'The majority of us accept the concept that we need to master simple skills before moving on to more complex'. Yet for many students, the challenge is precisely that one skill is not 'mastered' before another is taught. The Activity Based Learning (ABL) approach in India is based on the principle of scaffolding. ABL teachers give students regular

individual feedback on their learning and, as knowledge is acquired, it is recorded as segments of a snake. With modular lessons available, and frequent assessments, students who are absent, or slower to learn, can catch up from where they left off (Aslam and Rawal, 2016).

The challenge and costs to produce contextualized assessments that respond to the needs of children, parents, teachers, and administrators are significant. Can this approach be replicated elsewhere and what are the costs? Who were the critical stakeholders and what incentives were required to mobilize teacher development? What would the consequences be of putting the data in the hands of parents and children? Despite many questions, what this case study illustrates is the importance of international academics' close relationships with ACTRC and the political alignment that enabled data uptake.

References

Andrabi, T.; Das, J.; Khwaja, A. 2016. 'Report cards: the impact of providing school and child test scores on educational markets', Policy Research Working Paper no. 7226. Washington DC: World Bank.

Aslam, M.; Rawal, S. 2016. *Activity based learning, a multi-faceted investigation of a lauded reform: India and beyond, a synthesis report*. London: Department for International Development (DFID).

Banerjee, A.; Banerji, R.; Berry, J.; Duflo, E.; Kannan, H.; Mukherji, S.; Shotland, M.; Walton, M. 2016. *Mainstreaming an effective intervention: evidence from randomized evaluations of 'teaching at the right level' in India* (No. w22746). Washington DC: National Bureau of Economic Research.

Banerji, R. 2016. Mobilising community for better teaching and learning outcomes: field experiments in rural schools. In: M. Vipal (Ed.), *Claiming India from below: activism and democratic transformation*. London: Routledge.

Evans, D.; Popova, A. 2016. 'What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews', Policy Research Working Paper no. 7203. Washington DC: World Bank.

UK DFID (Department for International Development). 2011. *Participatory poverty and vulnerability assessment: understanding the regional dynamics of poverty with particular focus on Ghana's Northern, upper East and upper West regions.* London: DFID.

Commentary

Jorge Sequeira

Developing assessment assumptions that mirror our belief that a larger proportion of the population can learn more advanced skills is a good step forward. It reflects current trends on how to monitor the progress of SDG 4 to 'ensure inclusive and equitable quality education and promote lifelong learning opportunities for all', a key objective of the 17 SDGs that configure the 2030 Sustainable Development Agenda.

This agenda is integrated and places education at its core. Moreover, education has been given a standalone 2030 objective, and is also importantly present in the other SDGs. The response of the education sector to this challenge is through *lifelong learning*, i.e. the contribution of education in achieving the other SDGs (e.g. poverty, inequality, employment), and the required assessment(s) of learning outcomes to that end.

The assessment challenge is therefore twofold: a first tier of assessments that prepare students to learn more advanced skills (learning to learn, ICTs, soft skills, etc.) and a second tier that considers the education areas that are key to meeting the other SDGs (lifelong learning). In Care's example in the Philippines, the importance of a baseline per student and, especially, a clear understanding of 'behaviours' and their classification, respond to this endeavour. They will likely have a positive impact on the teaching/learning process, on teachers' overall performance, and on identifying the factors associated with learning.

However, the above-mentioned challenge poses an additional difficulty, as the curriculum, teachers, and expected outcomes for lifelong learning are commonly not managed by the national education system. This situation is more acute in the global south, where education services are commonly carried out by different ministries and entities. Managing assessment data and the capacity to link it to policy-making remain a concern.

More recently, the availability of international large-scale data on learning outcomes (TERCE, PISA, TIMSS, etc.) has generated much interest. Yet, the use of their results for evidence-based policy-making remains limited due to insufficient capacity to link results with policies, particularly at the BoP. Many countries still prefer one-off research

findings even though these studies are incomplete or of no relevance to the issue at stake.[28]

Assessments of learning outcomes are powerful instruments to advance education, provided there is institutional commitment to use their results. The implementation of SDG 4 and the education components of the other SDGs require such instruments, with students and learners at the centre of the teaching/learning process, such as is presented in the case of the Philippines.

28. Often countries develop assessments that consider cognitive outcomes (mathematics, language, reading, etc.) and ignore other dimensions e.g. relevance, efficiency, equity. See: www.unesco.org/new/en/santiago/press-room/newsletters/newsletter-laboratory-for-assessment-of-the-quality-of-education-llece/n16/08/

Chapter 10
Learning at the bottom of the pyramid in youth and adulthood: a focus on sub-Saharan Africa

Moses Oketch

Introduction

A good-quality education is now considered a human right and a global public good because of the economic and non-economic benefits that are associated with education. Today, education is viewed as a means to shared prosperity and an end in its own right, enabling individuals to fulfil their own potential and contribute to open societies. A recent report in *The Lancet* has projected that the Republic of Korea will exceed 90 years in average life expectancy by 2030, in large part due to its inclusive quality education that has promoted shared prosperity and enabled individuals to fulfil their potential (Kontis *et al.*, 2017). Yet, among the poor and marginalized in low-income countries, too few young people make the transition from primary to post-primary learning, and as a result many have reduced life-chances. If priority needs to be given to supporting this transition, what kinds of opportunities can build relevant life and labour skills and support civic participation for the marginalized group at the bottom of the pyramid in learning? What types of measurement tools can or should be used (or not used) in order to determine effective learning and effective policies for enhancing educational achievement at the BoP in the domain of youth and beyond?

Following the attention on the UN SDGs, many countries are poised to revive education policies around learning outcomes and their measurement. While examinations have been used to capture learning outcomes at the primary level, they have been insufficient in providing an evidence base for supporting inclusive learning for marginalized youth and adults who face special challenges in making the transition to post-primary learning. This is because high-stakes examinations, which are dominant in low-income countries in sub-Saharan Africa, are structured in such a way that they produce winners and losers – and poor, marginalized children often are the losers. They end up as youth at the BoP who have failed to make the post-primary learning

transition because of education systems in which such progress is based purely on meritocracy.

Non-formal education (NFE) could be a remedy if taken seriously at the level of national policy formulation and resourcing, and it is noteworthy that there is a resurgence of interest in this topic after nearly 40 years of debate and neglect (Britto, Oketch, and Weisner, 2014). Learning, however, is now the focus of such interest, which resonates well with the present attention on learning outcomes at the global level (Aspin *et al.*, 2001; Hanushek and Woessmann, 2008). If anything, there is already impetus given to non-formal learning through the SDGs framework, and a rise in technologies extending to those at the BoP. The reference to non-formal learning in the SDGs and in technological advances represents recognition that to reach and address the challenges of learning at the BoP in the domain of youth and adulthood, new approaches to learning will have to extend beyond formal systems.

This chapter will consider the roles of non-formal learning, with reference to the reality that formal school systems alone will not be able to cope with the challenges youth and adults at the BoP face when trying to transition into post-primary learning. Without alternative programmes for learning at the BoP in the domain of youth and adults, it is likely that a negative demographic transition will happen – in a few years' time, those at the BoP who have failed to make the transition to post-primary learning will become adults who lack relevant life and labour skills and civic engagement awareness, thereby extending their marginalization from youth into adulthood.

Given school systems' focus on meritocracy, it is no wonder that youth hold an unenviable position (Resnick and Thurlow, 2015), since many drop out of school or face limited intake or transition into post-primary learning. The concern is that on the one hand in such contexts, the youth who make the transition are considered 'agents of change' who are driven by the aspiration for a better life through their contributions to a productive labour force. They are portrayed as a 'youth dividend'. On the other hand, those at the BoP – because they have dropped out of school – are viewed as 'a lost generation' who are trapped by their BoP status and economic vulnerability. These youth and adults at the BoP can be found in the slums of cities where they seek highly vulnerable employment opportunities, and others are found idling in rural areas with little to do after failing to transition into post-primary learning. Yet, there is also agreement that acquiring skills relevant to current and future employment opportunities would

be key to leveraging the potential that African youth constitute for economic development and prosperity (Filmer and Fox, 2014).

Context for youth and adults at the BoP in sub-Saharan Africa

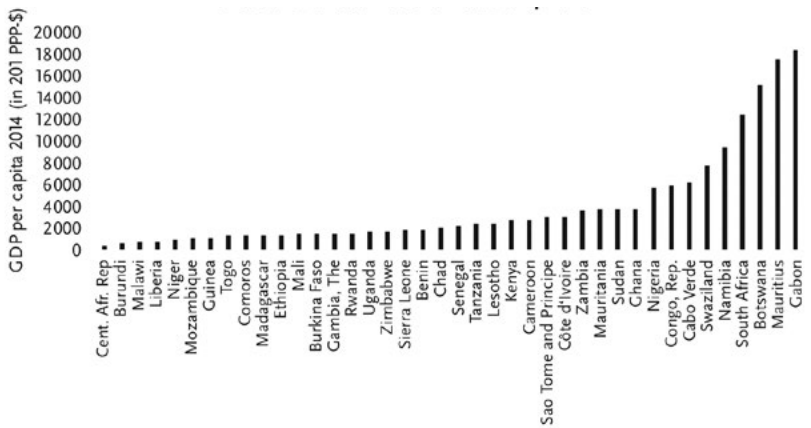
Economic trends across sub-Saharan African (SSA) countries have tremendous variability (see *Figure 10.1*). This has implications for analysing the challenges of learning at the BoP in the domain of youth and adulthood. A positive economic growth outlook is necessary to absorb skills into the labour market. Skills are in turn a necessary condition for a positive economic growth outlook, especially in systems that have in place strategies for addressing inequalities resulting from education systems. It can be argued that there is a feedback effect between the economy and skills development. Yet, the economic picture in SSA has not been systematically analysed through the lens of challenges of learning at the BoP, also associated with limited transition into post-primary education.

As shown in *Figure 10.1*, there are extremes in the economic growth pattern in SSA. The Central African Republic shows a very low GDP per capita of less than $1,000. In contrast, Gabon has the highest GDP per capita of about $20,000. However, in terms of challenges of learning at the BoP for youth and adulthood, at these two extremes they are likely to be similar. This is in part because these growth figures are driven by commodity goods such as oil in Gabon, and much insight could be drawn from them to advance our understanding of the challenges of learning at BoP in the domain of youth and adulthood. For many SSA countries, these economic indicators draw much attention, far more than the challenges of learning at BoP. Whereas, if these two were combined – such that interest in these economic indicators was matched with interest in learning, and specifically challenges of learning at the BoP in the domain of youth and adulthood – a much better policy evidence base might be generated that would serve to articulate better ways of addressing the learning challenges. So, unless there is attention to learning at the BoP for youth and adulthood, positive change will be very difficult.

In SSA, it is also clear that there is a youth 'bubble'. Some term this, optimistically, a 'demographic dividend', but a large investment in schooling is needed as well as sufficient labour market demand to leverage this demographic potential (Oketch, 2017). There are also large youth bulges, though in some countries these have started to decline (see *Figure 10.2*). This is a sign that peaks may have been

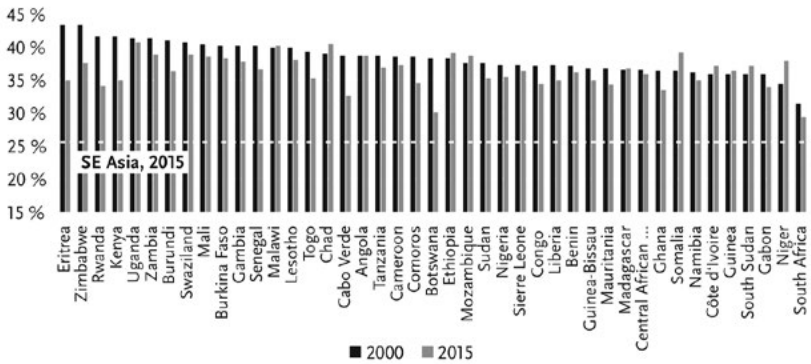
reached and a demographic dividend might happen – but these are general trends that do not address the challenges of those at the BoP, due in particular to the persistent problems of school dropout during and after primary schooling.

Figure 10.1 GDP per capita in sub-Saharan Africa



Source: Oketch, 2017.

Figure 10.2 Change in youth population aged 15–24 years between 2000 and 2015

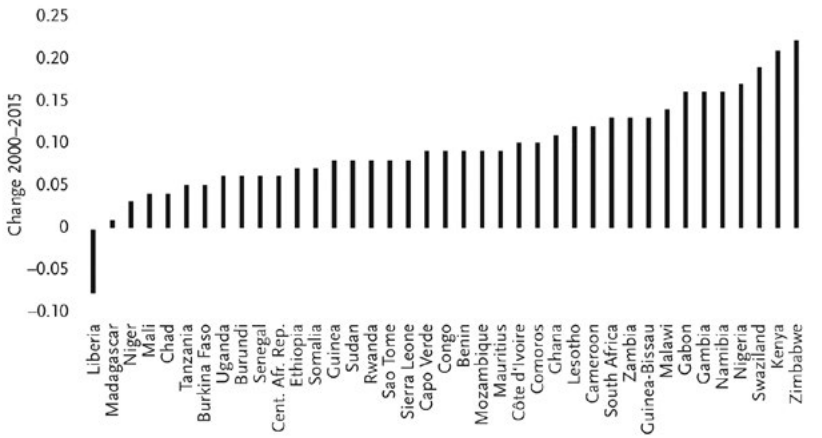


Source: Oketch, 2017.

Furthermore, low-quality schooling is associated with poor teaching methods and overcrowded classrooms, due in part to past high fertility rates, which have caused a major strain on education systems

in SSA. However, populations in Africa have higher rates of education today than ever before (see *Figure 10.3*). The figure displays the change in the fraction of the population 15+ with at least secondary education attainment. Apart from a drop in Liberia, secondary education has expanded from one percentage point in Madagascar to over 20 percentage points in Kenya and Zimbabwe (Oketch, 2017). This positive trend excludes the youth at the BoP. In addition, it only tells us about enrolment, and not learning outcomes. Assessing the learning of those at the BoP is becoming an ever more necessary endeavour.

Figure 10.3 Change in fraction of population 15+ with at least secondary education attainment



Source: Oketch, 2017.

To sum up, while there has been a generally positive economic trend in SSA in the past decade, there has not been an analysis on how this has impacted the youth at the BoP, and how it might be related to the challenges of learning at the BoP in the domain of youth and adulthood. Only when governments begin to understand and address the challenges of learning at the BoP in the domain of youth and adulthood will these indicators be more useful in realizing the 'demographic dividend.' For far too long, learning at the BoP has been left out of discussions concerning technical and vocational education and training (TVET), a topic to which we now turn.

The role of TVET

A World Bank (Filmer and Fox, 2014) report states that 11 million youth are expected to enter the labour market each year in SSA, and this trend will continue over the next decade. The majority of those who will fail to secure employment will be those at the BoP. They will also lack requisite skills to improve the quality of the available work in the informal sector, where many of them currently work. In this context, TVET is often proposed as a potential magic solution, and of strategic importance in addressing the special challenges faced by marginalized youth in Africa (African Union, 2007).

Yet, research on TVET and youth employability in SSA remains contested and uncertain (Oketch, 2007, 2015, 2017), often lacking a strong empirical base and analytical robustness (Fox and Thomas, 2016). Therefore, the relative effectiveness of TVET participation in improving learning and the labour market outcomes for young people at the BoP remains uncertain. Analyses that look at TVET participation, the degree to which it affects the post-primary learning transition, the wide range of effects on different individuals, and the stability of the potential effects over the working life would highlight the extent to which those at the BoP are impacted. This can be done by leveraging cross-national micro-level data to illuminate the effectiveness of TVET in different education systems and under varying labour market circumstances. This would surely advance our understanding of how the youth and adult learning transition and labour market interaction can be assessed under different contexts, to begin to understand and address the challenges of learning at the BoP.

In addition, qualitative system-level data on TVET policies and practices can be applied to better associate learning at the BoP with youth employment outcomes across countries. It is only after such analysis has been done that the role of TVET in this domain may be clearly assessed. At the moment, many countries have placed their hopes on TVET without a framework for how it can address the challenges of learning at the BoP. The focus should be on understanding how TVET systems address low levels of literacy and numeracy, limited functional language development, low attainment qualifications, weak social capabilities, and so on – all necessary employability component skills for youth at the BoP. Answering these questions requires an understanding of some specific TVET experiences.

TVET experiences in sub-Saharan Africa

There is a large body of literature on TVET in SSA straddling the disciplinary fields of education and economics. Much of it consists of single-country policy analysis, studies of access to TVET, and the changing nature of supply and demand more generally from a policy perspective. Biavaschi *et al.* (2012) provide an extensive review of this literature. Prominent within it is the mismatch between TVET provision and the labour market skills needed. Other significant trends are the low level of provision, the growing importance of private providers, and the role of new technologies. In most SSA countries, TVET has played only a marginal role to date (UK DFID, 2007; Oketch, 2007, 2017) despite recurring policy recommendations. Enrolment in vocational education as a share of all enrolled in secondary education has been noted to be below 10 per cent in most SSA countries – some exceptions include Liberia and Mali (with both reaching more than 30 per cent), as well as Angola, Rwanda, and Sierra Leone with more than 10 per cent (UK DFID, 2007; Atchoarena and Delluc, 2001).

In terms of unemployment rates and youth unemployment, the literature indicates that the position of young people in these labour markets is aggravated by a lack of education and training (Rioust de Largentaye, 2009; Garcia and Fares, 2008) and suffers from long transition periods from school to their first job, lasting between one year (Côte d'Ivoire) and 6.7 years (Mozambique) (Garcia and Fares, 2008). Other scholars have focused on the difficulties with the promotion of TVET that are attributed to issues such as its mismatch with young people's aspirations (Oketch, 2007; Atchoarena and Delluc, 2001; Foster, 1965). For extensive discussion see Biavaschi *et al.* (2012).

However, other evidence has pointed to the improved labour market performance of recent TVET graduates (Denu, Tekeste, and van der Deijl, 2005), while others have argued that young people turn mostly to self-employment or work unrelated to the skills learned (Lahire, Johanson, and Wilkox, 2011). Better skill formation and inclusion in the labour market, in particular due to the systematic involvement of the private sector, have also been reported (Rioust de Largentaye, 2009). Especially for marginalized youth and adults at the BoP with little formal schooling, apprenticeships may offer an avenue to improved labour market outcomes (Monk, Sandefur, and Teal, 2008). In contrast, some have argued that the development of functioning work-based training systems requires the contribution of the social partners, employers, and trade unions (Rioust de Largentaye, 2009). Yet this is often a difficult task, as the flexible,

informal forces of traditional apprenticeship can easily be distorted and overburdened by dependency on a supply-driven training programme (Mansuri and Rao, 2004). Palmer (2009) discusses several attempts of the Ghanaian government to formalize informal apprenticeships, and points to potential unintended ramifications in Ghana and other African countries.

In the end, TVET retains a difficult standing in SSA countries, despite some evidence that it can lead to better integration into wage employment, as demonstrated by the study on Ethiopia (Garcia and Fares, 2008; Guarcello, Lyon, and Rosati, 2008). Its current relevance in addressing the challenges of learning at the BoP for youth and adults remains debatable and unclear.

Gaps in evidence for improving learning at the BoP

First, there is need for more clarity on who the marginalized youth and adults are. Large populations in SSA have limited opportunities for transition into post-primary education, and there is even less systematic information and data available on those at the BoP and their special learning challenges. Having clear information, classification, and data of those at the BoP is a necessary step towards formulating a clear agenda for instituting programmes, such as NFE, that can address the special challenges of learning for youth and adults at the BoP.

Second, we need a better understanding of the challenges of learning at the BoP, in combination with the challenges or barriers of transitioning into post-primary education. This is necessary in order to develop appropriate programmes to address the needs of those at the BoP. This may include understanding and addressing barriers associated with education systems, such as meritocratic intake into limited spaces in secondary education, expansion of the education system, and making better-quality schools that are comprehensive and inclusive. NFE approaches need to be systematically developed and integrated with the knowledge and skills that those at the BoP already possess and utilize in their everyday economic and civic undertaking.

Among the issues that require attention is assessment of what BoP youth and adults know and are able to do (their skills). We need to know more about what they know, and how they utilize their knowledge specifically and broadly to enhance their informal learning. Tapping into the skills that they already utilize in non-formal learning would support their literacy, numeracy, and general economic and civic knowledge,

and impact their life chances. Obviously, the starting point is to ask how the skills of those at the BoP left out of the school systems can be measured. Once this is systematized and agreed upon, then skills training programmes can be developed and instituted.

These are some attempts to address these issues and offer clarity on the learning for this domain at the BoP (drawn from Gal, 2016: 5):

1. In the Literacy Assessment and Monitoring Programme (LAMP, 2009), the focus was on numeracy skills which are 'measured using short tasks with mathematical content that are embedded in hypothetical context that stimulate real-life situations'.
2. In the International Adult Literacy Survey (IALS, 1996), the focus was on quantitative literacy in relation to the knowledge and skills necessary to perform everyday mathematical tasks, such as figuring out a tip or determining the amount of interest on a loan or 'locat[ing] and us[ing] information contained in various formats (including job applications, payroll forms, transportation schedules, maps, tables and graphics)'.
3. In the Adult Literacy and Life Skills Survey (ALLS, 2005), attention was paid to skills that enable individuals to respond to 'mathematical demands of diverse situations', and 'numerate behaviour' is said to have occurred when 'people manage a situation or solve a problem in a real context'.

However, none of the studies cited above focused on the specific challenges of learning for those at the BoP. When developing policies for addressing the challenges of learning at the BoP for youth and adulthood, it will be necessary to determine the extent to which TVET can support a transition to post-primary learning, and to base such analysis on cross-country evidence about TVET uptake and labour outcome opportunities for those at the BoP. Subsequently, existing programmes would need to be reviewed to assess whether there are barriers or enabling conditions to facilitate the role of TVET in supporting learning for those at BoP. Integrating TVET and NFE can offer new learning platforms for understanding and addressing such challenges in SSA countries.

Conclusion

Youth and adult learning at the BoP in SSA will remain one of the most difficult challenges for meeting the 2030 UN SDGs. Clearly, the population of low-skilled youth is growing (even if more African children are going to school than ever before). The quality of learning is

known to be inadequate, leading to a growing number of poorly skilled youth and adults. While many SSA economies are growing, most of the growth is in the sectors that require at least secondary education, a goal that remains out of reach for many people throughout SSA.

In this chapter, TVET was discussed as one important avenue to address challenges of learning at the BoP in the domain of youth and adulthood, but it has also been noted that TVET is itself beset by several challenges. In the decades ahead, the majority of the population in many SSA countries will be at the BoP, including low-skilled youth and adults. At the same time, systematic programmes and attention to adult education are also lacking. There is also little research devoted to adult education and learning. Thus, an alternative focus that addresses the gaps in formal TVET and lack of attention to NFE will be needed. New approaches should include flexible and creative educational pathways and occupational training schemes geared for those at the BoP, and these should be an integral part of future policy (see Gal, 2017).

Economic indicators related to market growth or (un)employment tend to mask trends and problems related to those at the BoP, so how then would it be known if learning at the BoP has improved in the years since 2015? It is important to note that Goal 4.6 of the SDGs calls on countries to 'ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy' by 2030. Promoting but also monitoring Goal 4.6 will present multiple challenges to stakeholders. To address SDG 4.6, there is greater need for (a) further conceptualization of target skills of literacy and numeracy, as well as (b) improved indicators that can provide comparative data (see Gal, 2016).

Formal TVET options will not be able to adequately serve youth at the BoP until the quality and relevance of primary education equips them to continue successfully in secondary schools (see Muskin, this volume). Therefore, systems and their donors and other partners must look seriously to develop non-formal and informal learning and training options. In addition, the informal sector, which currently serves as the place to gain rudimentary trade skills, is another area of relevance for youth and adult learning at the BoP because of its prevalence in SSA among those who have not gained post-primary schooling (see Muskin, 2009).

In sum, given the rapidly changing demographics and economies across Africa, significantly more attention will need to be paid to the role of learning among youth and adults at the BoP. Non-formal learning is one important avenue to address this need, and especially to produce

skills that will enhance the employment and further education for African youth and adults at the BoP.

References

African Union. 2007. 'Strategy to revitalise technical and vocational education and training (TVET) in Africa'. Report to the Meeting of the Bureau of the Conference of the African Ministers of Education of the African Union, Addis Ababa, Ethiopia, 29–31 May.

Aspin, D.N.; Chapman, J.D.; Hatton, M.; Sawano, Y. (Eds). 2001. *International handbook of lifelong learning*. Dordrecht, Netherlands: Springer.

Atchoarena, D.; Delluc, A.M. 2001. *Revisiting technical and vocational education and training in sub-Saharan Africa.* Paris: IIEP-UNESCO.

Biavaschi, C.; Eichhorst, W.; Giulietti, C.; Kendzia, M.J.; Muravyev, A.; Pieters, J.; Rodríguez-Planas, N.; Schmidl, R.; Zimmermann, K.F. 2012. 'Youth unemployment and vocational training', Discussion Paper no. 6890. Bonn, Germany: Institute of Labor Economics (IZA).

Britto, P.R.; Oketch, M.; Weisner, T.S. 2014. 'Non-formal education and learning'. In: D. A. Wagner (Ed.), *Learning and education in developing countries: research and policy for the Post-2015 UN Development Goals* (pp. 74–90). Basingstoke, UK and New York: Palgrave Macmillan.

Denu, B.; Tekeste, A; Van der Deijl, H. 2005. *Characteristics and determinants of youth unemployment, underemployment and inadequate employment in Ethiopia*. Geneva: International Labour Office.

Filmer, D.; Fox, L. 2014. *Youth employment in sub-Saharan Africa*. Washington DC: World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/16608/9781464801075.pdf?sequence=1&isAllowed=y

Foster, P.J. 1965. 'The vocational school fallacy in development'. In: B. Anderson (Ed.), *Education and economic development.* Chicago, Ill.: Aldine.

Fox, L.; Thomas, A., 2016. 'Africa's got work to do: a diagnostic of youth employment challenges in sub-Saharan Africa'. In: *Journal of African Economies*, *25*(1), i16–i36.

Gal, I. 2016. 'Assessment of adult numeracy skills', background paper prepared for the 2016 Global Education Monitoring Report. Paris: UNESCO.

——. 2017. Comments on draft paper 'What are the special challenges of learning at the BoP in the domain of youth, young adulthood, and adulthood?' Learning at the Bottom of the Pyramid Conference, University of Pennsylvania, March.

Garcia, M.; Fares, J. (Eds). 2008. *Youth in Africa's labor market*. Washington DC: World Bank. http://documents.worldbank.org/curated/en/687001468007775569/pdf/454880PUB0Box311OFFICIA0L0USE0ONLY1.pdf

Guarcello, L; Lyon, S.; Rosati, F. 2008. 'Child labor and youth employment in Ethiopia'. In: M. Garcia and J. Fares (Eds), *Youth in Africa's labor market* (pp. 181–223). Washington DC: World Bank.

Hanushek, E.A.; Woessmann, L. 2008. 'The role of cognitive skills in economic development'. In: *Journal of Economic Literature, 46*(3), 607–668.

Hanushek, E.A.; Schwerdt, G.; Woessmann, L.; Zhang, L., 2016. 'General education, vocational education, and labor-market outcomes over the life-cycle'. In: *Journal of Human Resources, 51*(2), 49–88.

Kontis, V.; Bennett, J.E.; Mathers, C.D.; Li, G.; Foreman, K.; Ezzat, M. 2017. 'Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble'. In: *Lancet, 389*(10076), 1323–1335.

Lahire, N.; Johanson, R.; Wilcox, R. T. 2011. *Youth employment and skills development in the Gambia*. Washington DC: World Bank.

Mansuri, G.; Rao, V. 2004. 'Community-Based and -Driven Development: A Critical Review'. In: *The World Bank research Observer, 19* (1), 1-39. http://documents.worldbank.org/curated/en/178951468336565202/pdf/764740JRN0Comm0Box0374379B00PUBLIC0.pdf

Monk, C.; Sandefur, J.; Teal, F. T. 2008. 'Does doing an apprenticeship pay off? Evidence from Ghana', CSAE Working Paper Series 2008–08. Oxford, UK: Centre for the Study of African Economies, University of Oxford.

Muskin, J. 2009. 'Tinkering with the tinker: meeting training needs in the informal sector of Chad'. In: R. Maclean and D. Wilson (Eds), *International handbook of education for the world of work, bridging academic and vocational learning*. New York: Springer.

Oketch, M.O. 2007. 'To vocationalise or not to vocationalise? Perspectives on current trends and issues in technical and vocational education and training (TVET) in Africa'. In: *International Journal of Educational Development*, *27*(2), 220–234.

——. 2015. 'Education policy, vocational training, and the youth in sub-Saharan Africa'. In: J. Thurlow and D. Resnick (Eds), *African youth and the persistence of marginalization: employment, politics, and prospects for change* (pp. 133–151). London and New York: Routledge.

——. 2017. 'Cross-country comparison of TVET systems, practices and policies, and employability of youth in sub-Saharan Africa.' In: F. Eicker, G. Haseloff, and B. Lennartz (Eds), *Vocational education and training in sub-Saharan Africa: current situation and development* (pp. 25–38). Bielefeld, Germany: Bertelsmann.

Palmer, R. 2009: 'Formalizing the informal: Ghana's national apprenticeship programme'. In: *Journal of Vocational Education and Training*, *61*(1), 67–83.

Resnick, D.; Thurlow, J. (Eds). 2015. *African youth and the persistence of marginalization: employment, politics, and prospects for change*. New York: Routledge.

Rioust de Largentaye, A., 2009. 'Vocational training and the informal economy'. In: *MENA development report: unlocking the employment potential in the Middle East and North Africa*. Washington DC: World Bank.

UK DFID (Department for International Development). 2007. 'Technical and vocational skills development', briefing paper. London: DFID.

Commentary

Iddo Gal

Oketch's paper reviews many important issues related to learning at the BoP in SSA, mainly from an economic or human capital perspective. My comments focus on two key topics, of relevance to low-income countries, both in SSA and elsewhere: (a) equal attention to adult education and (b) greater attention to teaching and monitoring of numeracy and mathematics.

First, concerning equal attention to adult education. The paper highlights gaps in the transition from primary to secondary education and to the employment market, and questions the educational options available for out-of-school youth. As Oketch notes, while these are critical targets to address, he acknowledges that changes in education *quality* (vs. mere access) at the primary and secondary levels are slow, and its improvement will face many obstacles. Hence, I argue that a forward-looking educational policy, and programmes for workforce integration, should not focus so intensively on youth, as Oketch does, since in coming decades the majority of the adult population in many high BoP countries will continue to have (relatively) low skills, even if skills of the youngest cohorts of adults (slowly) improve.

I thus believe that countries with high BoP populations should not ignore the skills of adults (15+ to elderly) who are beyond the formal school years, since their competencies will affect economic growth and the characteristics of the job markets into which we hope to see integration of school graduates. In order to ensure economic and social progress, future educational policy and research should encompass flexible and creative educational pathways, learning opportunities, and technical and occupational training schemes, geared for adults as well as for youth at the BoP.

Second, concerning more attention to teaching and monitoring of numeracy and mathematics. Oketch argues that general economic indicators related to labour market growth or (un)employment may mask trends and problems related to those at the BoP. If we agree with that, how then would we know if learning at the BoP has improved, in the post-2015 education agenda? Goal 4.6 of the UN SDGs calls on countries to 'ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy' by 2030. Promoting but also monitoring Goal 4.6 presents multiple challenges to stakeholders.

Efforts to conceptualize and measure basic skills of adults in low- and middle-income countries have traditionally focused on literacy (i.e. language skills). While literacy is critical, policy and field-level responses are separate for *numeracy* and literacy. Further, 'numeracy' and 'mathematical knowledge' are related but not the same, yet both are a key gateway to productive employment in many occupations and to technological progress, and for managing personal affairs and creating an empowered citizenship. To address Goal 4.6, we need (see Gal, 2016): (a) further conceptualization of target skills, (b) better indicators that shed light on systemic factors related to numeracy and mathematics learning (e.g. curriculum coverage, teachers' qualifications), and (c) better monitoring tools that can provide comparative data on levels of numeracy skills both of adults and of school-age learners, whether they are in or out of school.

References

Gal, I. 2016. 'Assessment of adult numeracy skills'. Background paper commissioned for the UNESCO Global Education Monitoring Report 2016: Education for people and planet. Paris: UNESCO. Retrieved from: http://unesdoc.unesco.org/images/0024/002455/245573E.pdf

Commentary

Joshua Muskin

Moses Oketch makes an important contribution to the discussion of meeting learning needs at the BoP, exploring particularly the education and training needs of youth, young adults, and adults. Exposing the enormity of this challenge in SSA, he rightly links the problem to the endemic failures of systems to provide adequate quality post-primary education opportunities to BoP-dwellers. Importantly, he suggests further that meeting the challenge is a matter not simply of rights, but also of meeting nations' economic and social goals.

Raising important questions about these dimensions, he points to a penury of research by which to delve more deeply into the core query of his title. Looking to the TVET sector as the most promising pathway to raising youth's prospects for emerging from their economic and social marginalization, Oketch finds little to describe how this happens in reality. TVET is clearly a good place to start, but in many countries, it will continue to fall well short of its promise until at least three conditions are met. First, as Oketch notes, the supply of TVET in terms of skills and skilled youth must better match the demand. Second, the scope of accepted modalities for its delivery must expand greatly, including the informal apprenticeship system as well as training in core competencies, such as literacy. Lastly, TVET will not serve youth adequately at the BoP until the quality and *relevance* of primary education equips them to continue successfully in secondary education or training.

Looking at the second point, systems and their donor and other partners must look seriously at developing non-formal and informal learning and training options. This does not exclude TVET.[29] It does, however, prioritize meeting youth, young adults, and adults where they are – professionally, personally, physically, and in terms of their cognitive skills – to help them gain and continue to strengthen the literacy, numeracy, 'life', and professional skills they require to have more secure, fulfilling lives.

Strategically, a wide range of learning options should be offered, coupled with efforts to strengthen the overall formal education and training system, which will be slower to gestate. Oketch alludes to one

29. My organization, Geneva Global, is experimenting with a promising strategy in Ethiopia that brings secondary school dropouts back into secondary school after many years away by offering them weekend TVET courses.

of these options: the informal sector. The informal sector is currently a place to gain rudimentary trade skills, but there is great potential for those in the BoP to benefit in terms of both their cognitive and technical skills, as well as to develop behaviours and methods of entrepreneurship (Muskin, 2009). The many functional literacy courses offered by governments and NGOs may (and often do) extend beyond basic competencies to offer lifelong learning options across a range of relevant life topics. And the realm of ICT promises ever-increasing access to information that individuals at the BoP can use to elevate their knowledge, their productivity, and their fulfilment.

Formal education and training paths to escape the BoP are surely important; but more immediate gains may be achieved by promoting non-formal and informal learning options that aim first to raise the BoP floor.

References

Muskin, J.A. 2009. 'Tinkering with the tinker: meeting training needs in the informal sector of Chad'. In: R. Maclean and D. Wilson (Eds), *International handbook of education for the world of work: bridging academic and vocational learning* (pp. 245–60). Dordrecht, Netherlands: Springer.

Chapter 11

'Learning at the bottom of the pyramid' and the global targets in education

Aaron Benavot

Introduction

This paper first examines key notions in a 'learning at the bottom of the pyramid' approach, and the extent to which they reflect or capture recently adopted global targets on education (under SDG 4) as part of the 2030 Agenda for Sustainable Development. It then discusses the complexities involved in estimating and analysing disparities in education among marginalized learners, groups, and populations, which are relevant for SDG 4 targets and for identifying low-achieving learners. It goes on to argue that much more conceptual and empirical work needs to be completed before a 'learning at the bottom' approach can effectively contribute to the ways in which governments address, through policy and practice, equity-related challenges in education. At the very least, future work would need to concentrate on three key questions:

- Which children and youth are being left behind, thereby impeding the achievement of universal completion of a full cycle of primary and secondary education?
- Which policies and practices are most likely to enhance the provision of good-quality education and training, and result in the acquisition of a broad array of skills and relevant learning outcomes for children, youth, and adults?
- Which factors and forces greatly impede the ability of marginalized groups to gain access to relevant lifelong learning opportunities for life and for work?

Initial distinctions

Each key concept in a 'learning at the BoP' approach deserves further scrutiny. Learning, in particular, should not be considered synonymous with knowledge, or the proficiencies and skills acquired in school-based programmes and courses. Learning begins at birth and continues throughout life. Young children learn informally in non-school settings

through observation, imitation, trial and error, and myriad other experiences, which result from the actions of parents, caregivers, and others who populate their environment (see *Chapter 1*). While much early learning is informal or incidental in nature, the settings in which learning occurs, and the purposes for which learning is directed, become increasingly formalized and instrumental as children age.

Formal education settings structure learning through a sequence of institutionalized categories (Benavot, 1997). Nested instruction in the early grades typically focuses on language acquisition, basic numeracy, and the arts; in the upper grades instruction includes content from the natural and social sciences and focuses on higher-order thinking skills, as well as the inculcation of relevant cultural content (valued knowledge, norms, attitudes, behavioural expectations, and worldviews). Increasingly, colleges and universities are emphasizing course offerings and degree programmes deemed to be more relevant to the demands of the labour market. Many schools and higher education institutions are also promoting individually tailored learning platforms that enable self-directed learning. Moreover, non-formal education and training occur in planned learning settings (e.g. centres, associations, online, libraries, workplaces), which are outside the formal system. Non-formal learning activities are often job-related, but also provide training in life skills and in other forms of personal growth. *Table 11.1* presents a schematic view of the myriad settings, programmes, and frameworks in which learning occurs over the life course.

By and large, researchers who probe the notion of 'learning at the BoP' tend to privilege a rather limited range of the many settings in which learning takes place over the life cycle. Whether by design or by default, research in this area mainly examines tested knowledge and skills taught in formal settings during the years of compulsory schooling. As I argue below (see also Wagner *et al.*, Introduction to this volume), not only is this conception of learning problematic, it also fails to capture the many forms of learning embedded in the SDG 4 targets (and in the targets of other SDGs that have an explicit education dimension).

To the extent that we seek to understand patterns of learning among those at the lower rungs of some metric – in other words, those 'at the bottom' – two approaches can be discerned. The first approach ranks learners according to indicators of their socio-economic or socio-demographic status, such as wealth, poverty, ethnicity, linguistic or religious minority status, parents' educational background, migration status, or cultural capital. Information on several of these measures can also be combined into an index to rank learners. Such an approach to

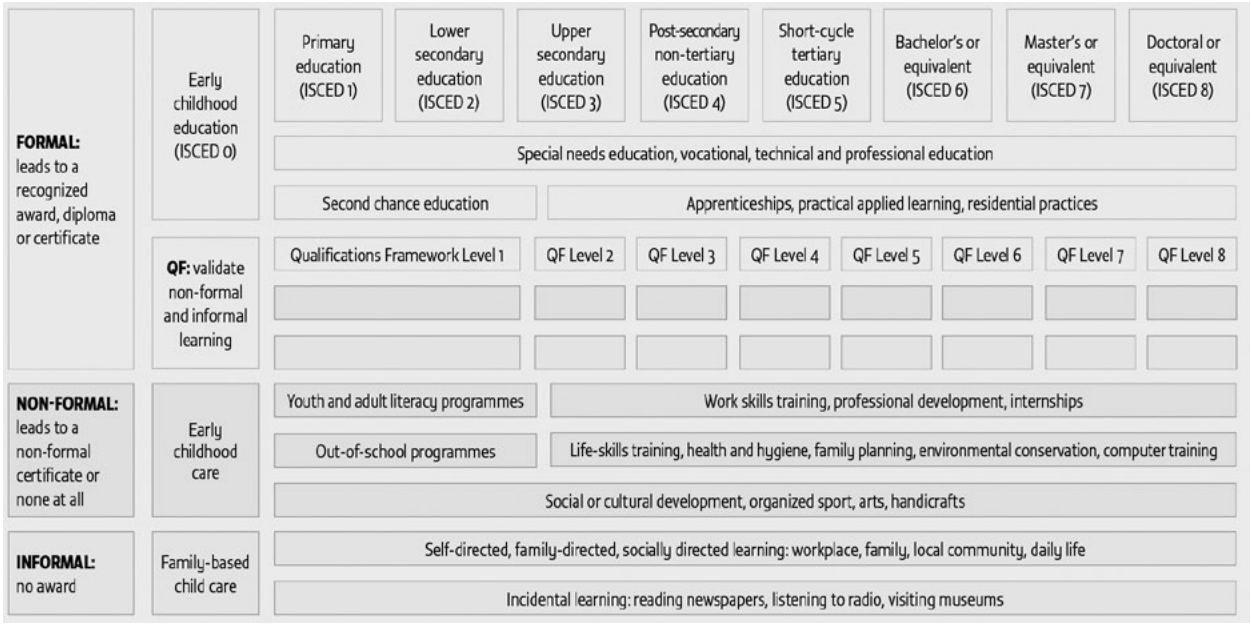
'learning at the bottom' attempts to identify the strengths and weaknesses of those students belonging to the lowest ranked socio-demographic groups in each context, and then highlight efforts or policies found to be most effective in raising group learning levels.

The second approach focuses on identifying underachievers according to their performance on some summative learning assessment, regardless of their SES. This approach highlights the strengths and weaknesses of poor performers, as well as the potential steps or interventions that could be taken to address their specific learning needs. The latter might include, for example, smaller classes, more gender- or culturally sensitive textbooks, a nutritious daily meal, instructional materials in relevant languages, supplemental tutoring, or accelerated learning programmes.[30] Education researchers often refer to such learners as 'low achievers' or 'underachievers.' Meta-analyses of policies and interventions thought to improve learning levels among low achievers are fairly extensive, and more so in higher-income than in lower-income countries. Given the association between weak performance and lower SES, policies found to improve learning levels are often – though not always – thought to influence opportunities for improved SES.

Perhaps the most curious term in the phrase 'learning at the bottom of the pyramid' is the word 'pyramid' itself. The typical argument is that if one plots the distribution of income of the world's population, or that of a particular country, then one finds this to be in the shape of a pyramid. Low-income countries (defined in terms of per capita GDP or GNP) or people living in poor households (defined in terms of household income falling below the official poverty line of a country) form the wide base of the pyramid, which narrows as either country or household income levels increase (see *Chapter 2*). In fact, the actual shape of the distribution depends to a considerable extent on: (a) the economic indicator used; (b) whether or not one weights it by population (thereby underscoring the importance of income growth in China and India); and (c) whether one conceives of well-being as a single-dimensional (economic) status or a multi-dimensional

30. I would argue that this approach is less applicable to adult learners, since there are few regularly conducted standardized summative assessments of adult skills and proficiencies, according to which low achievers could be identified. Moreover there is no standardized information about the many non-formal (and sometimes formal) educational settings in which they may have participated to acquire or augment their skills.

Table 11.1 Settings in which learning occurs over life



Source: UNESCO, 2016.

one involving economic, social, and political components.[31] If the idea is simply to highlight the special challenges found in the education systems of low-income countries or schools attended by poor people, then perhaps the notion of the pyramid may obscure more than it reveals.

Furthermore, if we focus our attention on the distribution of learning, based on a particular assessment, in order to identify low achievers, as noted above, then the distribution of learning in a given population rarely looks like a pyramid, but rather a bell curve, at times with a truncated left tail or a slightly elongated right tail. With these considerations in mind, this paper focuses on issues related to the learning outcomes of children, youth, and adults at the lower ends of a learning distribution, in a wide range of contexts (both low and high income), rather than on the notion of a 'pyramid'.

The importance of equity in global education policy is increasingly recognized

The focus on improving the prospects of learners populating the lower rungs of some scale, whether defined according to learners' SES or socio-demographic status or by their educational performance, is clearly aligned with the emphasis on equity in the recently adopted 2030 Agenda for Sustainable Development (UN, 2015). Indeed, learning and equity are at the heart of the fourth global goal on education (SDG 4): 'Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.' Most of the 10 SDG 4 targets focus implicitly or explicitly on aspects of quality education and equity, with special attention paid to reducing disparities between dominant or advantaged populations, on the one hand, and vulnerable or marginalized ones, on the other.

Equity issues also animate the discourse and consensus statements of the international education community. In May 2015 the World Education Forum in Incheon, Republic of Korea brought together over 1,600 governmental and non-governmental participants, who adopted an SDG-aligned platform called Education 2030, with a strong focus on equity. The final Incheon Declaration calls inclusion and equity 'the cornerstone of a transformative education agenda' and commits governments to develop and implement education policies that address

31. Traditional income measures are based on an estimate of the total value of goods and services flowing through a market and are criticized for minimizing the value of goods or services traded in kind, or produced for household subsistence, and for undervaluing the contributions of women to household subsistence or economic progress.

'all forms of exclusion and marginalization, disparities and inequalities in access, participation and learning outcomes'. It makes this point emphatic by stating: 'no education target should be considered met unless met by all,' and focuses on the disadvantaged and marginalized 'to ensure that no one is left behind'.[32]

Arguably, interest in and attention to equity in the formulation of international education targets have increased in recent decades. From Jomtien and through Dakar, ensuring universal access to basic education and the provision of good-quality education were the two key principles undergirding global education policy. In previous decades, much less attention was given to inequalities in education beyond reducing the number of out-of-school children and gender disparities in access and completion. One contributing factor to the increasing importance of equality in education is that governments and policy analysts have far more comparative evidence on disparities in education and learning outcomes today than they did in the past. Indeed, the number of data sources on these topics has grown dramatically, as has the circulation of prominent publications reporting on education inequalities across and within countries. Coverage of these issues has also increased in the traditional media, as well as on social media. While media outlets and commentators often focus on a small set of high-performing countries – e.g. Japan, the Republic of Korea, Singapore, Finland, and even Cuba – and the policies they have put into place to improve student performance, some also discuss equity issues. For example, commentators make reference to specific policies that: (a) target underachieving students, (b) prioritize the reduction of learning disparities, and (c) create a more inclusive education system. Many analysts see equity-oriented or pro-poor policies as valuable, not only because they improve the overall pool of knowledge and skills among future workers, but also because they increase a country's competitiveness in a knowledge- and innovation-driven global economy.

In reality, inequalities in education are ubiquitous and enduring

Inequalities in access, retention, and measurable learning outcomes are rampant in national education systems. While the extent of these inequalities may vary by education level, stage of learning, and by the particular disadvantaged group(s) found at the bottom of a specific indicator, inequalities in education and learning are ubiquitous. Moreover,

32. http://en.unesco.org/world-education-forum-2015/incheon-declaration

educational inequalities are long-lasting: despite extensive policy borrowing, waves of reform, and 'proven' policy interventions, they tend to persist over time and across generations, especially among children, youth, and adults from marginalized communities.

There are many reasons that inequalities in education endure. They include, but are not limited to, the following:

- Most vulnerable groups have limited power and agency to affect policy reform in areas that most impact their lives.
- Many decision-makers believe, incorrectly, that policies found to be effective in addressing the challenges faced by 'average' or typical learners will be equally effective in addressing those faced by learners from marginalized groups.
- Many policies targeting disadvantaged communities are improperly or only partially implemented.
- Mobilizing funds and resources to specifically address the needs of the disadvantaged is difficult.

The influence of these forces tends to be especially salient in the Global South. The larger point is clear: notwithstanding international declarations and policy priorities, most inequalities in education are deeply rooted, and not easily overcome. Evidence of the many – and often subtle – ways education conveys or reinforces disadvantage from one generation to the next deserves greater attention (Antoninis, Delprato, and Benavot, 2016). Indeed, it is precisely this knowledge gap that serves as an underlying motivation for the current volume.

Assessing the impact of markers of disadvantage on education

Social markers of disadvantage are used to identify inequalities in education. Many such markers are common across diverse countries, cities, and communities, while others are more context-specific. Among the most common and widely used markers are low SES, gender, and rural residence. The impact of these factors, when combined, becomes even more powerful. Poor, rural girls and women are substantially worse off on almost all educational measures than their wealthy, urban male counterparts. That said, many context-specific markers of disadvantage (not necessarily mutually exclusive) negatively impact an array of diverse groups – for example, racial, ethnic or linguistic minorities; persons with disabilities; migrants, refugees, and internally displaced persons; urban slum dwellers; pastoralists and nomads; and street children and orphans. Members of these groups encounter special obstacles with

respect to education. Estimating the extent of educational disparities among these groups is often difficult, if not impossible, given the lack of relevant disaggregated data and the limitations of sampling strategies currently employed in household surveys or government censuses. Thus, some of the most powerful effects of disadvantage on education are 'invisible' and go largely unmeasured, at least in quantitative terms.

In recent decades, our understanding of the impact of disadvantage on learning outcomes has grown notably with the proliferation of international, regional, national, and citizen-led learning assessments (Kamens and Benavot, 2011; Benavot and Köseleci, 2015). These assessments provide a wealth of disaggregated data on learning outcomes for different populations, broken down by both common and less common social markers. Detailed, relevant information about particular disadvantaged groups is more likely to be found in assessments organized by national authorities or national NGOs, given their awareness of the historical disadvantages faced by certain communities or social groups. This is true with one important exception. National assessments, if they are school-based, as most are, cannot provide information about two key marginalized groups: those who have never gained access to school, and those who have left school for whatever reason and are no longer enrolled. As such, school-based assessment platforms tell us very little, if anything, about the learning levels of these two groups and the challenges they face.33 Only through learning assessments carried out in households do we gain some measure of the scale of educational disadvantage experienced by all members of an age cohort.

To be clear: the backgrounds of and obstacles faced by learners at the bottom of distributions are diverse and often context-specific. They do not easily fit into existing categories or respond to existing policy solutions. To the extent that pedagogical strategies and programme interventions fail to capture the incredibly diverse backgrounds and needs of students whom they target, they are unlikely to succeed. Evidence as to which school- and home-based interventions are most effective – including second-chance programmes, non-formal education, training and professional development programmes, and policies targeting the most marginalized – is surprisingly limited and inconclusive (see *Chapter 10*; Snilstveit *et al.*, 2015). The value of the

33. Many low achievers in a given learning assessment are 'at risk' of dropping out. Thus, information about them gleaned from the assessment can tell us something about the social, psychological, and pedagogical challenges they face.

policy implications they render remains modest. More often than not, the research designs of such studies are driven by the agendas of donor partners and international aid agencies, who seek to determine the effectiveness of their investments, rather than which programmes or reforms are best designed to serve the needs of the marginalized.

What kinds of learning are invoked in the SDG 4 targets?

One of the main innovations of the new Sustainable Development Agenda was to place learning and learning outcome indicators at the core of its education goal (SDG 4). The ways in which learning is embedded in the formulation of SDG 4 targets, and the nature of global indicators approved by the UN to assess country progress on SDG 4, reflect the comprehensive and universal nature of the 2030 Agenda. They are also unprecedented in the range and diversity of learning they invoke (*see Table 11.2*).[34] They include learning that enables toddlers and young children to be prepared for primary school, minimum proficiencies in reading and mathematics to be achieved by primary and lower secondary students, foundational skills in literacy and numeracy, and useful TVET and employability skills for youth and adults, as well as various types of knowledge and skills related to sustainable development, global citizenship, and other key concepts that all learners should possess.

Of the 10 SDG 4 targets, learning issues are central in five of them. They include targets 4.1 (primary and secondary students achieving relevant and effective learning outcomes), 4.2 (early children's readiness for primary education), 4.4 (relevant skills for employment, decent jobs, and entrepreneurship), 4.6 (achieving adult literacy and numeracy), and 4.7 (all learners acquiring knowledge and skills needed for sustainable development and global citizenship).[35] In addition, a closer inspection of the global indicators for all 10 SDG 4 targets shows that learning-related measures are not only prevalent in the five aforementioned targets, but also in two others: indicator 4.3.1 (which looks at the participation rate of youth and adults in formal and non-formal education and training in the previous 12 months)

34. The MDGs made no mention of learning-related outcomes in the education goal. Of the six Education for All (EFA) goals, two mentioned learning-related outcomes: EFA goal 4 focused on reducing adult illiteracy by half and EFA goal 6 specified that literacy, numeracy, and life skills should be improved through formal education, although this did not result in the generation of comparable global data on learning outcomes, particularly beyond primary school.

35. A more in-depth analysis of the monitoring and measurement challenges related to targets 4.4, 4.6 and 4.7 can be found in Benavot and McWilliam, 2016, Benavot and Koseleci, 2016, and Benavot and Lockhart, 2016.

and indicator 4.a.1 (which highlights the importance of access to computers and the internet for pedagogical purposes).

Thus, a country truly committed to the 2030 education agenda would need to seriously (re)consider the purposes of its education system, the ways in which teachers are prepared, the contents conveyed by textbooks and emphasized in assessments, and the ways the system monitors and evaluates what students actually take away from their educational experiences. Such a country would also need to rethink how out-of-school children, youth, and adults gain access to education and learning opportunities, including professional development and training, in a lifelong learning perspective (Benavot, 2017).

Box 5.3. SDG 4 Targets and Global Indicators (learning issues are *italicized*)

4.1 By 2030, ensure that all girls and boys complete free, equitable, and quality primary and secondary education leading to *relevant and effective learning outcomes.*

4.2 By 2030, ensure that all girls and boys have access to quality early childhood development, care, and pre-primary education so that they are *ready for primary education*.

4.3 By 2030, ensure equal access for all women and men to affordable quality technical, vocational, and tertiary education, including university.

4.4 By 2030, substantially increase the number of youth and adults who have *relevant skills, including technical and vocational skills, for employment, decent jobs, and entrepreneurship.*

4.5 By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples, and children in vulnerable situations.

4.6 By 2030, ensure that all youth and a substantial proportion of adults, both men and women, *achieve literacy and numeracy.*

4.7 By 2030, ensure all learners *acquire knowledge and skills needed to promote sustainable development,* including among others through *education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship, and appreciation of cultural diversity and of culture's contribution to sustainable development.*

4.a By 2030, build and upgrade education facilities that are child, disability, and gender sensitive and provide *safe, non-violent, inclusive, and effective learning environments* for all.

4.b By 2020, substantially expand globally the number of *scholarships* available to developing countries, in particular least developed countries, SIDS and African countries, for enrolment in *higher education, including vocational training and information and communications technology, technical, engineering and scientific programmes,* in developed countries and other developing countries.

4.c By 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing states.

Global Indicators

4.1.1 *Proportion of children and young people (a) in Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education achieving at least a minimum proficiency level in (i) reading and (ii) mathematics,* by sex.

4.2.1 Proportion of children under 5 years of age who are *developmentally on track in health, learning and psychosocial well-being,* by sex.

4.2.2 *Participation rate in organized learning* (one year before the official primary entry age), by sex.

4.3.1 *Participation rate of youth and adults in formal and non-formal education and training* in the previous 12 months, by sex.

4.4.1 Proportion of youth and adults with *information and communications technology (ICT) skills,* by type of skill.

4.5.1 Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated.

4.6.1 Percentage of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex.

4.7.1 Extent to which (i) global citizenship education and (ii) education for sustainable development, including gender equality and human rights, are mainstreamed at all levels in (a) national education policies, (b) curricula, (c) teacher education, and *(d) student assessment.*

4.a.1 Proportion of schools with access to: (a) electricity; (b) *Internet for pedagogical purposes; and (c) computers for pedagogical purposes.*

A learning at the bottom approach and the SDG 4 targets

In principle, a learning at the bottom approach *can* contribute to progress in the new global agenda in education. Essentially, this would mean focusing on the particular kinds of learning challenges faced by low-income countries, on the one hand, and, on the other, by children, youth, and adults from poor families or marginalized communities in all countries. In doing so, it would not only speak to the equity orientation of the SDGs, but potentially help pinpoint policies, programmes, and practices that have been shown to bring about significant improvement in learning outcomes, especially among low achievers and learners from poor households or marginalized communities. Too little is currently known about which targeted reforms and policy interventions make a lasting difference to disadvantaged learners. We also need to know about how equity-orientated teaching and learning practices enable disadvantaged learners to acquire knowledge, skills, values, and worldviews relevant to the SDGs.

To be sure, this is a tall order. And not easily accomplished given the current state of available data. Two of the many recommendations made by the 2016 UNESCO GEM Report would support activities related to a learning at the bottom approach. First is the call encouraging regional associations and networks to develop peer learning mechanisms and to exchange information on relevant SDG 4 issues – for example, tackling disadvantage in education, in-depth analyses of learning among the marginalized, or initiatives in the area of Education for Sustainable Development (ESD) and Global Citizenship Education (GCED) (UNESCO, 2017). Second is the call to establish a research hub for global education measurement issues, in order to find ways to reach greater consensus on measuring SDG 4 learning outcomes, and possibly support the construction of an international household survey programme dedicated to education, which could help fill in existing indicator and measurement gaps.

However, the larger issue is whether supporters of a learning at the bottom approach will go beyond easily accessible measures of learning – namely, school-based surveys of a narrow range of learning outcomes at the primary and lower secondary level – and engage with the broader and more comprehensive learning agenda proposed by SDG 4. To the extent that researchers seriously interrogate data related to the knowledge and skills acquired by children, youth, and adults in multiple formats and settings, especially those from disadvantaged and marginalized backgrounds, they can seriously contribute to debates about what is needed to transform education systems, and ensure greater inclusiveness and equity. To the extent they remain wedded to conventional analytic strategies, their contribution is likely to be muted.

There is a growing body of evidence that the effective inclusion of learners from marginalized groups can mitigate economic disadvantage, end poverty, reduce inequality, and foster economic growth. Based on such evidence, the 2016 GEM Report (UNESCO, 2016) argued that countries should invest in making their education systems more inclusive and equitable, which would bring about more sustainable, just, and peaceful societies. An expanded learning at the bottom strategy has the potential to provide a much-needed platform to critically assess the current policies that promote inclusive education.

Conclusion

This paper shows that while learning and equity are deeply embedded in the recently adopted international education targets, our existing knowledge base is limited in several respects. Most studies focus on a narrow range of learning outcomes in primary and secondary education, mainly in relation to basic proficiencies in reading and mathematics. Relevant learning outcomes in pre-primary frameworks, in TVET programmes, in higher education, and in a diverse array of adult non-formal education programmes are rarely studied, although they are integral to the SDG 4 vision. Existing data draw mainly on school-based surveys of learning, which are largely cross-sectional in design. Few studies follow learners as they move through different stages of the life cycle or provide information on the availability of and participation in lifelong-learning opportunities (e.g. job re-training, professional development, ICT training, self-directed learning, cultural knowledge). Information about some of the least-developed countries in the world and the most marginalized communities is often missing. Finally, we have yet to build a stockpile of robust evidence about how existing interventions and targeted approaches improve learning levels over time, especially among less-advantaged learners.

This list of the research challenges should give pause for reflection. It should also suggest to supporters of a learning at the bottom approach what needs to be done in order to make a significant contribution to current global commitments in education.

What is clear is that inclusive education systems that increase learning levels for all and reduce inequalities in education are critical for progress in SDG 4. We need a well-established knowledge base that enables us to respond to the following key question: *What improvements in the types of knowledge, skills, and other learning outcomes among low achievers – or learners from marginalized groups – are needed and in which types of contexts, in order to contribute to progress in SDG 4 targets?* Answers to this question would provide a convincing rationale for reforms and strategies that target substantive improvements among marginalized learners located at the bottom of learning scales.

In sum, while the growing recognition of the importance of equity in lifelong education is welcome, there is an urgent need for greater clarity of definitions, consistency in data collection and measurement, and a more deliberate approach to building evidence on how best to improve relevant learning outcomes for disadvantaged children, youth, and adults in different learning contexts. In order to close persistent

disparities among subgroups within countries, we must consolidate what we know, ask questions about what we don't, develop the right tools and metrics, and carry out policy and programmatic research on effective solutions. We need to pay particular attention to policies and practices that are most conducive to improving education inclusion and reducing inequality. Research focused on strategies to improve 'learning at the bottom' can expand and improve our knowledge base, as well as its relevance to the ambitious SDG education agenda.

References

Antoninis, M.; Delprato, M.; Benavot, A. 2016. 'Measuring inequality in education'. In: International Social Science Council (ISSC), Institute of Development Studies (IDS), and UNESCO, *Challenging inequalities: pathways to a just world, World Social Science Report 2016* (pp. 63–67). Paris: UNESCO.

Benavot, A. 1997. 'Institutional approach to the study of education'. In: *International encyclopedia of the sociology of education* (pp. 340–45). Oxford, UK: Elsevier Science.

——. 2017. 'Education systems need a major rethink'. In: *Ministers reference book, Commonwealth 2017* (pp. 69–72). London: Henley Media Group.

Benavot, A.; Köseleci, N. 2015. 'Seeking quality in education: the growth of national learning assessments, 1990–2013', paper commissioned for EFA Global Monitoring Report 2015. Paris: UNESCO.

——. 2016. 'Preparing learners for a different future: monitoring education for sustainable development and global citizenship in the new education agenda'. In: *SangSaeng* (Fall), 8–13.

Benavot, A.; Lockhart, A.S. 2016. 'Monitoring the education of youth and adults: from EFA to SDG 4'. In: H. Hinzen and S. Schmitt (Eds), *Agenda 2030: education and lifelong learning in the Sustainable Development Goals*, International Perspectives in Adult Education 75 (pp. 55–63). Bonn, Germany: DVV International.

Benavot, A.; McWilliam, A. 2016. 'Monitoring competencies and skills for life and work'. In: *Adult Education and Development*, 83, 44–49.

Kamens, D.; Benavot, A. 2011. 'National, regional and international learning assessments: trends among developing countries, 1960–2009'. In: *Globalisation, Societies and Education, 9*(2), 285–300.

Snilstveit, B.; Stevenson, J.; Phillips, D.; Vojtkova, M.; Gallagher, E.; Schmidt, T.; Jobse, H.; Geelen, M.; Pastorello, M.; Eyers, J. 2015. 'Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review, 3ie final review'. London: International Initiative for Impact Evaluation.

UNESCO. 2016. *Education for people and planet: creating sustainable futures for all*, 2016 Global Education Monitoring Report. Paris: UNESCO.

——. 2017. 'Strengthening peer learning of education policies for SDG 4: the role of regional organizations', Global Education Monitoring Report technical paper. Paris: UNESCO.

United Nations. 2015. *Transforming our world: the 2030 Agenda for Sustainable Development*. New York: United Nations.

Commentary

Nicholas Burnett

Benavot does four very useful things. He unpacks alternative interpretations of the BoP, noting that this can be income-determined or learning-determined. He underlines that equity is increasingly recognized in global education policy, even as income inequality within countries continues to rise. He reminds us how little is known about how to improve educational performance among different poor and disadvantaged children. Finally, he calls for 'research focused on strategies to improve "learning at the bottom"' that 'can expand and improve our knowledge base as well as its relevance to the 2030 UN SDG education agenda'.

These are all excellent points. They don't go far enough, however, because they are restricted to a research recommendation and do not consider if a BoP approach could also help prioritize within SDG 4, and lead towards any key indicators. And even that recommendation is only to do research to improve learning at the bottom; it is not even to prioritize such research.

Benavot uses the list of the SDG 4 targets and global indicators to show how learning permeates the goal. The table's very length, however, also demonstrates dramatically just how many elements there are in SDG 4, which covers all levels of education. This comprehensiveness is both a strength and a weakness of SDG 4. It is a strength that, unlike the preceding Education for All goals which ran through 2015, there is recognition of the importance of all parts of the education system. But it is a weakness, again unlike the EFA goals, that there is no prioritization among the various elements of SDG 4, since, realistically, no government can focus on all of them.

It is here that Benavot misses the opportunity to use a BoP approach to prioritize within SDG 4. Such an approach should focus us on those who are not and have not been learners, which would lead us back to prioritizing two of the former EFA goals in addition to a focus on learning within school: the 263 million children out of school and the 758 million adult illiterates (more like a billion, probably). This is not a minor point – SDG 4 and its attention to learning in school has led to a markedly reduced emphasis on out-of-school children and adult illiterates, the latter not ever being a high global priority.

In sum, a BoP approach could help to prioritize among SDG 4 indicators, of which there are currently too many, again with no prioritization. This should lead us to focus on disaggregated data and to pay most attention to the lower quintiles. We should also reconsider whether one simple indicator – equivalent to stunting for nutrition – could summarize goal progress, including equity. Possibilities are school life expectancy (even though it excludes learning), or the learning performance of girls in the bottom quintile or, as Wagner has proposed, a new Gini coefficient for learning (Wagner, 2018). Such an indicator would essentially say that an education system is only as good as how successfully it educates society's poor and disadvantaged.

Reference

Wagner, D. A. 2018. *Learning as development: rethinking international education in a changing world*. New York: Routledge.

Commentary

David K. Evans and *Fei Yuan*

In order to ensure learning among the most vulnerable children and youth, educators (from education policy-makers to classroom teachers) must know what helps those children and youth to learn. It seems straightforward. Yet recent years have seen a dramatic increase in what we know about improving learning. In the year 2000, we identified only 19 rigorous impact evaluations of interventions in low- and middle-income countries that reported learning outcomes.[36] By the year 2016, that number had risen to 300, a 15-fold increase (Evans and Popova, 2016; World Bank, 2017).

But as Benavot highlights, 'policies found to be effective in addressing the challenges facing "average" or typical learners' will not necessarily be effective in addressing those 'faced by learners from marginalized groups'. In this increased evidence on what works to improve learning, how much have we learned about improving learning for the most vulnerable? In a review of 281 evaluations with learning outcomes between 2000 and 2016, drawn from two databases (Evans and Popova, 2016; 3ie, 2016), we identified less than 11 per cent that reported outcomes separately for students with low SES. Only 23 per cent reported outcomes separately for students with initially low learning levels. Thirty-three per cent separated results by sex.

Significant resources are invested in these evaluations, and since this is the subset of evaluations with learning outcomes, most measure learning outcomes pre- and post-intervention. So the cost of identifying impacts differentially for the lowest performers should be low. Likewise, gathering data on sex and basic socio-economic indicators would cost little, but significantly deepen our understanding of which programmes are reaching the learners who need them most. For example, an evaluation in Chile that did a separate analysis by SES found that the 'quality of teacher–student interactions' was 'positively correlated with the performance of low-income students' (Bassi, Meghir, and Reynoso, 2016). A natural extension of this finding – meriting future research – is that investments in the quality of those interactions may particularly benefit the most

36. Rigorous impact evaluations are programme evaluations that include a credible identification of the 'counterfactual' (what would have happened in the absence of the programme), usually established with the use of a comparison group, through either a randomized controlled trial or a high-quality quasi-experimental design.

vulnerable learners. However, few evaluations take the SES of students into account. If research is to benefit the learners at the BoP, then it has to identify learning outcomes for those learners in particular.[37]

References

3ie (International Initiative for Impact Evaluation). 2016. 'Impact evaluation repository'. London: 3ie. Retrieved from: www.3ieimpact.org/en/evidence/impact-evaluations/impact-evaluation-repository/

Bassi, M.; Meghir, C.; Reynoso, A. 2016. 'Education quality and teaching practices', working paper 22719. Cambridge, Mass.: US National Bureau of Economic Research.

Evans, D.K.; Popova, A. 2016. 'What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews'. In: *World Bank Research Observer*, *31*(2), 242–270.

World Bank. 2017. *World development report 2018: learning to realize education's promise*. Washington DC: World Bank.

37. Researchers may be hesitant to report differentiated effects by SES, initial performance, or sex, because of concern over identifying spurious correlations that can emerge from testing many hypotheses. Two solutions to this concern include pre-specifying differentiated analysis at the outset of the study and – when not pre-specified – reporting such analysis as exploratory, but still reporting it.

Chapter 12
Is it sustainable to leave the bottom behind in the process of educational development?

Dirk Van Damme

Introduction

In September 2015, the world's leaders gathered in New York to set ambitious goals for the future of the global community. Goal 4 of the Sustainable Development Goals (SDGs) seeks to ensure 'inclusive and equitable quality education and promote lifelong learning opportunities for all'. More specific targets and indicators spell out what countries need to deliver by 2030. SDG 4 creates a formidable opportunity to advance the cause of providing quality education to all of the globe's children and adults.

Two aspects of Goal 4 distinguish it from the preceding Millennium Development Goals (MDGs) on education, which were in place between 2000 and 2015. Firstly, Goal 4 is truly global. The SDGs establish a universal agenda; they do not differentiate between rich and poor countries. Every single country is challenged to achieve the SDGs, including the relatively rich countries gathered in the OECD.

Secondly, Goal 4 puts the quality of education and learning outcomes front and centre. Access, participation, and enrolment, which were the main focus of the MDG agenda, are still important. The world is still far from providing equitable access to high-quality education for all. An estimated 57 million children still do not have access to primary education, and too many children continue to be excluded from the benefits of education because of poverty, gender, ethnicity, where they live, and armed conflicts. But participation in education is not an end in itself. What matters for people and for our economies are the skills acquired through education. It is the competencies and character qualities that are developed through schooling, rather than the qualifications and credentials gained, that make people successful and resilient in their professional and private lives. They are also key in determining individual well-being and the prosperity of societies.

One year later – and one year closer to the 2030 deadline – it is clear what the main challenges are. Of course, a main condition of success is the available resources that can be mobilized, and each country's capacity for implementation and reform. But behind that are two more serious challenges. The first one is developing a sound measurement agenda. Prior international experience in the realm of education, and also experience with other global issues such as climate change, has taught us that large-scale policy processes rely on a trustworthy and ambitious measurement agenda. For many of the targets in SDG 4, we simply lack reliable data, especially when it comes to measuring learning outcomes and skills. The Education SDG thus also provides a challenging agenda for creating tools that measure what really counts in education. The international research community cannot leave this agenda aside. The OECD is very engaged in this process, and willing to utilize all of its expertise to execute this agenda successfully.

The second challenge is probably even more critical. SDG 4 and its targets demand a huge amount of political will and enthusiasm. The fact that the world has agreed on these goals is already a huge success, and evidence of the shared understanding across the political spectrum of the importance of education for the future of humanity and this planet. Even with recent and still-to-come drastic shifts in the political and ideological hegemony in many parts of the world, I do not think this belief in the power of education will fade away. Why not? Simply because the stakes are too high. If we do not make all possible efforts to implement the education agenda successfully, prospects for a better world – however you wish to define it – look grim. The chances provided by the Education SDG are just too big to accept failure.

In this context, the concept of 'learning at the bottom of the pyramid' is a very useful one. It forces us to divert educational development away from national averages and the performance of the better-off and instead focus on those children and adults who have the fewest opportunities, and those parts of societies where challenges are biggest. Educational development will only be inclusive and sustainable if those at the bottom of the social and educational pyramid benefit from it, and if opportunities to benefit from learning are distributed in a more equitable way.

This paper will explore the scale of learning at the BoP, as measured by the share of people without a foundation level of educational attainment, and the share of young people without adequate foundation skills. We will focus on the inclusiveness of the distribution of educational attainment and learning outcomes. In doing so, we will make the case

for a concept of 'tolerance' of low educational opportunities and low relative learning outcomes, as measured by the length of the tail of the distribution. We will then examine whether this tolerance is a function of average levels of learning outcomes in societies or of overall social inequality. At the end of the paper we will argue for assessing relative learning exclusion as an autonomous variable, and make a case for an inclusive approach to educational progress and learning in the context of the SDG framework.

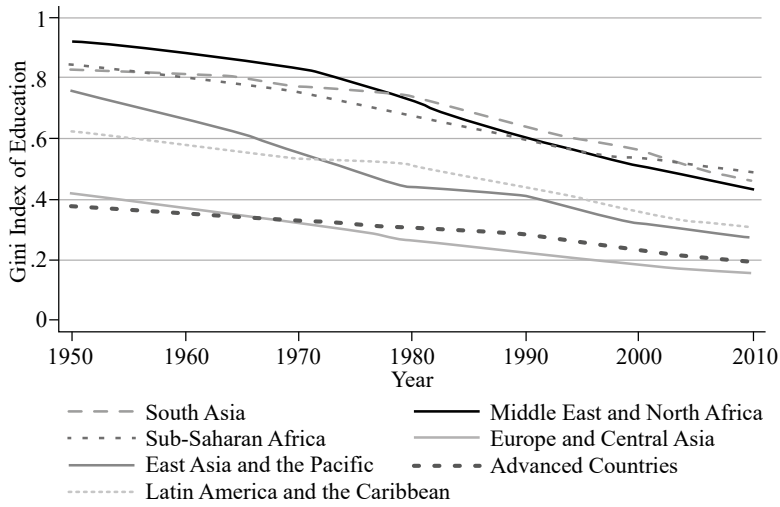
Educational expansion and inequality in educational attainment levels

Since the Second World War, the world has seen an unprecedented increase in educational participation and attainment. All levels of education have seen drastic growth, and global average years of schooling are much higher now than 50 years ago. Despite all these worldwide improvements, some countries have lagged behind, mainly in sub-Saharan Africa.

The expansion of education is visible in all indicators, especially average years of schooling and educational attainment levels (for learning outcomes, see next paragraph). The problem with such measures is that focusing on national or regional averages oversimplifies the realities within these areas. It is not always easy to focus the attention of statisticians and policy-makers to how the shape of the pyramid has changed as a consequence of educational expansion. Benaabdelaali, Hanchane, and Kamal (2011) usefully attempted to assess the inequality in educational attainment over time by applying the methodology of the Gini coefficient, most often used for measuring income inequality. *Figure 12.1* shows that inequalities in educational attainment have levelled out over the last 60 years, though this occurs at a different pace for each region. Still, in South Asia, Africa, and the Arab world, levels of inequality in educational attainment remain at an unacceptably high level. The Global Education Monitoring Reports extensively document the various indicators that measure the size of educational exclusion and disadvantage in the world.

The rapid expansion of educational opportunities over the past decades was most visible at the top of the distribution, that is, in the growing share of tertiary-educated adults. But educational opportunities also opened up at the bottom of the distribution and, as a result, the number of poorly educated people decreased. In other words, the entire distribution of educational attainment moved upwards. However, the

Figure 12.1 Education attainment Gini by region



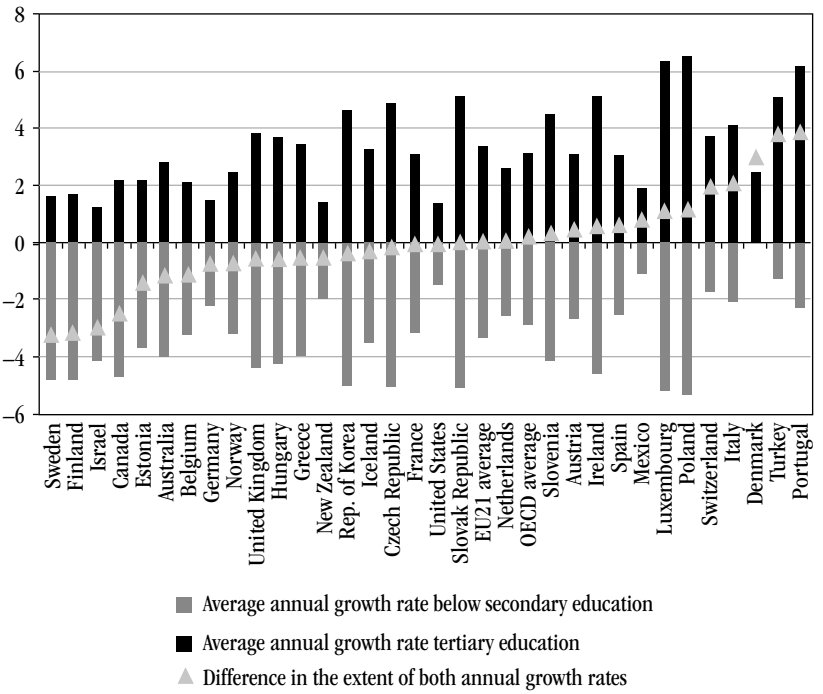
Source: Benaabdelaali, Hanchane, and Kamal, 2011.

speed of change can be different at the two ends of the attainment distribution. If the change at the top exceeds that of the bottom, then inequality in educational attainment increases. When people are left behind as access to education expands, social cohesion is threatened and inclusive growth is thwarted. There is ample evidence that educational exclusion comes with huge risks to health, employment, income, and even intangible qualities such as interpersonal trust, tolerance, and adherence to democratic values. A lack of educational opportunities also seems to be one of the main channels through which poverty and social inequality are transmitted from one generation to another.

By contrast, a process of inclusive growth, with equivalent growth at both ends of the spectrum, or the bottom end improving even faster, seems to be a good thing in itself. When societies become highly educated, routes towards many other opportunities in life become open.

A simple exercise based on educational attainment trend data for OECD countries between 2000 and 2013 shows that there are huge differences between countries in the average growth rates at the top and the bottom of the attainment distribution. *Figure 12.2*, based on calculations from data in the OECD Education Database, shows how the growth rates differ for tertiary education attainment and below-secondary

Figure 12.2 Average annual growth rates in below-upper-secondary and tertiary education, OECD countries, 2000–2013



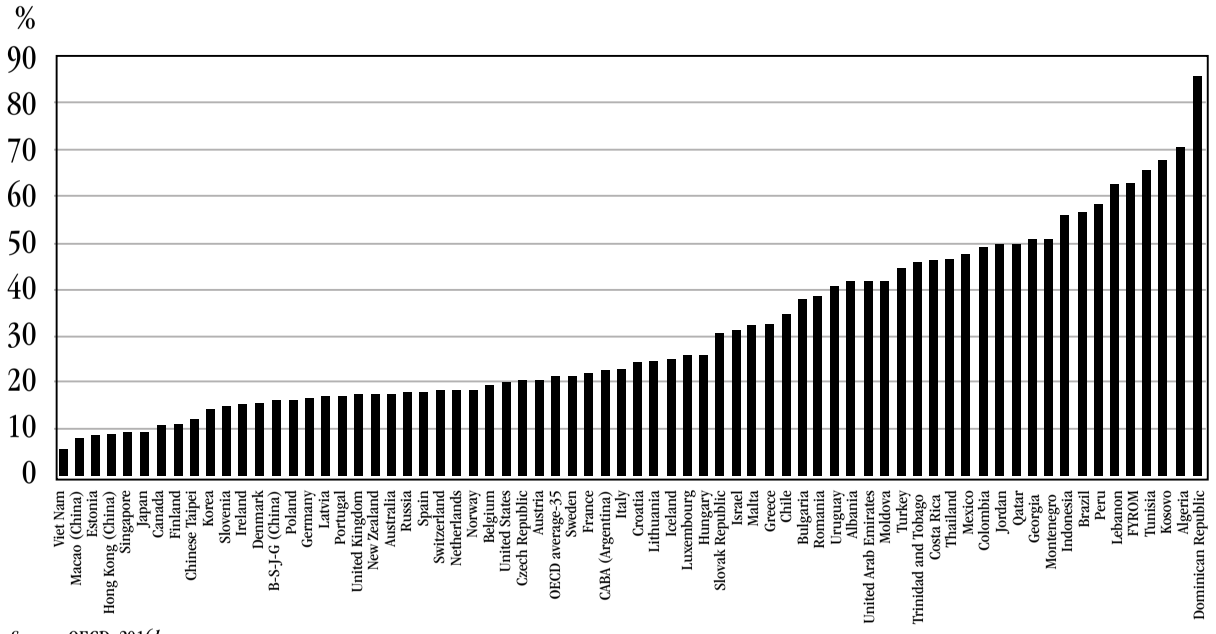
Source: OECD Education Database.

educational attainment. By calculating the difference between the two, we can build an indicator of inclusive growth of educational attainment.

Between 2000 and 2013 the share of tertiary-educated adults grew by 3.1 per cent per year on average in OECD countries, while the share of people without an upper secondary education decreased by 2.9 per cent per year on average. So, on average across OECD countries, the educational attainment distribution widened slightly.

But, as is clear in *Figure 12.2*, the differences among countries are huge. The chart shows the average annual growth rates at both ends of the distribution and compares the extent of both. At the left are Sweden, Finland, Israel, and Canada, where the average annual rate of reduction in the share of people without an upper secondary education was more than 2 percentage points greater than the average annual rate of increase in the share of tertiary-educated adults. Over this period, these countries prioritized reducing the number of poorly educated individuals over

Figure 12.3 Percentage of students below proficiency level 2 in PISA 2015 science assessment, PISA database, 2015



Source: OECD, 2016*b*.

Note: B-S-J-G (China) refers to the cities of Beijing, Shanghai, Jiangsu, and Guangdong in China. CABA (Argentina) refers to Ciudad Autonoma de Buenos Aires.

increasing the number of highly educated individuals, partly because they had already expanded the top end of the distribution. In these countries, the breadth of the distribution of educational attainment narrowed.

At the other end of the distribution are Portugal, Turkey, Italy, and Switzerland, where the average annual rate of increase in tertiary attainment was more than 2 percentage points greater than the average annual reduction in the share of people without an upper secondary education. In these countries, the distribution of educational attainment widened. Denmark is a special case because it is the only country in which the share of people without an upper secondary education increased between 2000 and 2013. Still, with increases at both ends of the spectrum, the distribution widened in Denmark too.

The total length of the two bars provides an indication of the overall growth in educational attainment. The greatest changes took place in the Czech Republic, Luxembourg, Poland, and the Slovak Republic, closely followed by Ireland and the Republic of Korea. In contrast, the overall change was smallest in Mexico, New Zealand, and the USA. But the size of overall change is unrelated to differences in the annual rate of growth at each end of the spectrum. This suggests that it is not the speed of change which determines whether the expansion of educational attainment is more or less inclusive. Rather, it is the policy environment around educational change that determines whether individuals at the bottom of the distribution also see their educational opportunities improve.

Countries that are in the process of becoming better-educated societies, where educational qualifications and skills determine income, well-being, and many other factors, should invest in improving opportunities across the population, not only for the most educated. With the right inclusive education policies in place, no part of the population risks being left behind and without access to the social and economic benefits that accrue to more educated people.

The bottom of the pyramid: the share of low performers

Of course, educational attainment only tells part of the story. When it comes to achieving the benefits of education, the quality of learning matters most. OECD's Programme on International Student Assessment (PISA) provides the world's largest and best collection of data on learning outcomes. Since its launch in 1997, PISA has become a leading reference on the quality of education systems worldwide. 'What is it important for citizens to know and be able to do?' is the question that underlies the PISA metric for quality, equity, and efficiency in school

education. PISA assesses the extent to which 15-year-old students have acquired the knowledge and skills that are essential for full participation in modern societies. An analysis of the PISA data allows us to assess the size of the bottom of the learning pyramid.

The latest PISA round, PISA 2015, focused on science, but also included assessments of reading (literacy) and mathematics (numeracy). Its results were published in December 2016. There were 72 countries participating, including a wide range of emerging and developing countries, among them Brazil, Indonesia, Viet Nam, Thailand, Jordan, the Dominican Republic, and Peru. The results of PISA 2015 are worrisome to the extent that they show how many young people fail to achieve even the most basic learning outcomes. *Figure 12.3* shows that in nearly two-thirds of the participating countries, over 20 per cent of 15-year-old students did not master the baseline proficiency level in science. Further, 21.2 per cent of 15-year-old students in OECD countries performed below Level 2, but this percentage varied from 8.8 per cent in Estonia and 9.6 per cent in Japan to 44.5 per cent in Turkey and 47.8 per cent in Mexico (OECD, 2016*b*). Among the non-OECD countries, the percentages varied from 5.9 per cent in Viet Nam to 70.8 per cent in Algeria and 85.7 per cent in the Dominican Republic. Looking at the very low part of the skills distribution, namely those students who failed to achieve even Level 1, the percentages are obviously much lower, ranging from a few tenths of a percentage in most OECD countries to 15.8 per cent in the Dominican Republic. Percentages can remain statistical artefacts; absolute numbers can bring life into them. In PISA 2012, the absolute number of 15-year-olds who performed below Level 2 in all three subjects was 1.2 million in Brazil, 1.7 million in Indonesia, 267,000 in Colombia, and 181,000 in Thailand.

SDG 4 seeks to ensure 'inclusive and equitable quality education and promote lifelong learning opportunities for all'. This requires that 'all learners acquire the knowledge and skills needed to promote sustainable development' (Target 4.7). Only in Canada, Estonia, Finland, Hong Kong (China), Japan, Macao (China), and Singapore do at least four out of five 15-year-old students master the baseline level of proficiency in all three domains of science, reading, and mathematics. This fact shows that there are countries on nearly every continent that could achieve the goal of universal basic skills by 2030. At the same time, the small group of countries that has moved close to securing basic skills for all shows how much remains to be done in most countries – including some of the wealthiest OECD countries – to attain the SDGs.

Figure 12.4 demonstrates that the share of low achievers in PISA (in this case for reading) has been declining in most countries between 2000 and 2012. In some countries, such as Peru and Indonesia, this decrease has been quite remarkable, with 20 and 13 percentage points respectively.

Obviously, the share of low achievers is largely a function of the overall performance of the education system in a country, but there are remarkable differences between countries with respect to the relationship between the PISA mean score and the share of low achievers. Countries such as Peru, Indonesia, Jordan, and Viet Nam have lower numbers of low achievers as could be predicted by their mean score, while countries such as Thailand, Turkey, and Serbia have a larger share of low achievers than predicted by their mean score (OECD, 2016*a*). This suggests that the tail of the skills distribution and the inclusiveness of educational performance vary quite a bit, and, hence, education policies matter in shaping the distribution.

The tail to the bottom of the pyramid: a measure of tolerance for exclusion in learning?

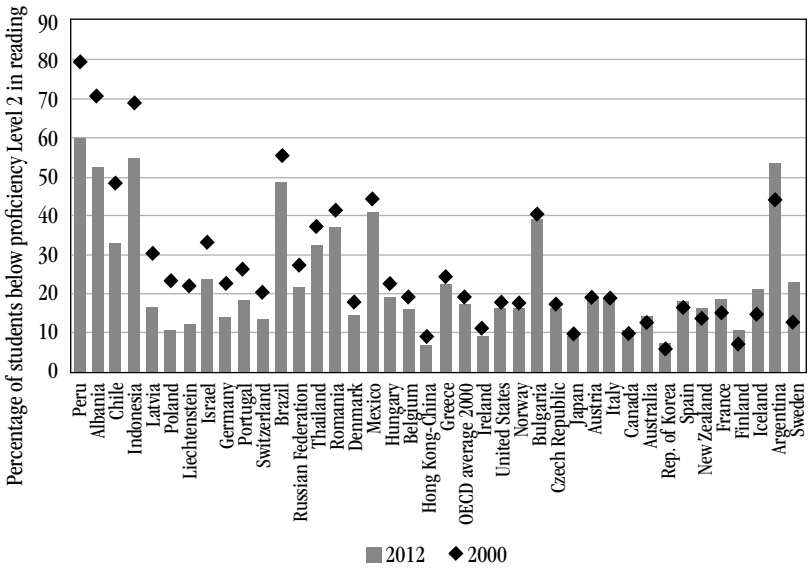
We will now explore this relationship in greater depth by looking at the length of the tail of the learning outcomes distribution. More specifically, we will take a look at the score point difference between the median and percentile 10. A large difference suggests that a country allows the bottom of the learning distribution to sink to a relatively low level. We could consider this difference an indication of tolerance for low learning achievement.

Across countries that participated in the PISA 2015 assessment, the score point difference between the median and percentile 10 varied from around 90 score points on the PISA scale to over 170. The interesting thing is that there is almost no correlation between the span between the median and percentile 10 and the average score of a country, as *Figure 12.5* demonstrates. In fact, the length of the tail of the learning outcomes distribution is, on average, slightly higher among countries with a relatively high average score. Some well-performing countries allow the bottom to sink quite low, while some of the less well-performing countries have a very compressed tail of the learning outcomes distribution. But there are also examples of the opposite.

Compare for example Costa Rica and Bulgaria, two countries with approximately the same average score on the PISA 2015 reading assessment, but with quite different lengths of the tail from the median to

the percentile 10 score. In Costa Rica the tolerance for low performance is much smaller than in Bulgaria. Or compare France with Viet Nam, two countries with similar average scores on the PISA scale, but Viet Nam's tail of the distribution does not go as low as France's.

Figure 12.4 Many countries have decreased their numbers of low-performing 15-year-olds, PISA database, reading, 2000, 2012

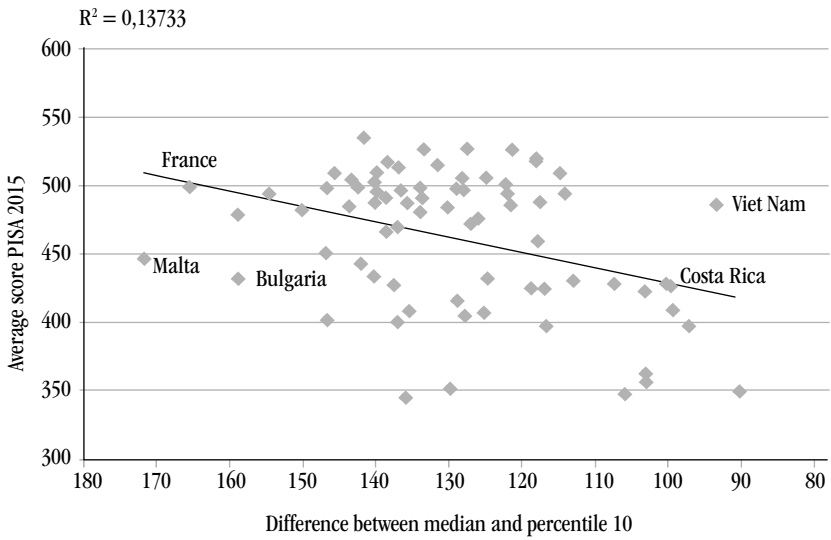


Source: OECD, PISA database.
Note: Countries are ranked in descending order of progress between 2000 and 2012.

It does seem that the tolerance for low performance is not a function of the average PISA score of a country, but the result of more or less deliberate policies to leave people at the bottom of the learning outcomes distribution behind. We thus come to the same conclusion as we did when looking at the inclusiveness of the growth of educational attainment.

Would the tolerance for low learning outcomes in a given society then be a function of overall social inequality? This hypothesis suggests that the bottom of the learning pyramid would be lower in societies where overall social inequality is higher. To test this hypothesis we can use another data set, namely the OECD Programme of International Assessment of Adult Competencies (PIAAC), which tested the foundational skills of 16–64 year-old adults in a number of countries.

Figure 12.5 The relationship between the average PISA 2015 score for reading and the length of the tail of the distribution (score point difference between median and percentile 10)



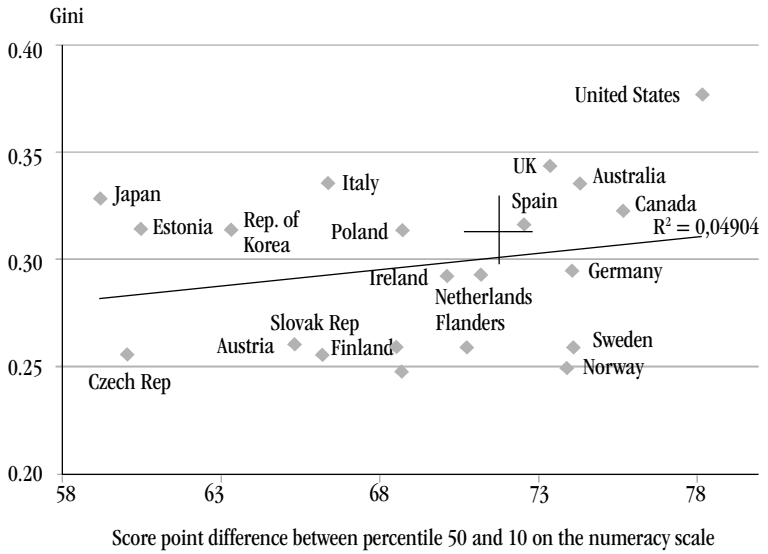
Source: PISA database, 2015.

In this case we compare the distribution of the numeracy skills. Although the share of adults with low numeracy skills is closely related to overall social inequality as measured by the Gini coefficient (countries that enjoy greater social equality, such as the Nordic countries and Flanders [Belgium], generally have fewer low-skilled and more high-skilled adults; correlation of .59), the length of the tail in the skills distribution itself is not related to overall social inequality, as is shown in *Figure 12.6* (Van Damme, 2014). Again, we come to the conclusion that the distance between the bottom and the median of the skills distribution is a societal artefact in its own right, independent from either the average level of skills in that society or from overall social inequality. We can thus view the tolerance for low learning outcomes as an independent feature of societies.

Learning at the bottom as a political issue

If the tolerance for relatively low learning opportunities and outcomes is seen as an autonomous feature of societies, relatively independent from the average level of skills in the country or the level of social inequality, then it seems legitimate to conclude that policies really matter.

Figure 12.6 Relationship between the length of the tail of the skills distribution among adults (score point difference between median and percentile 10) in numeracy and the Gini coefficient

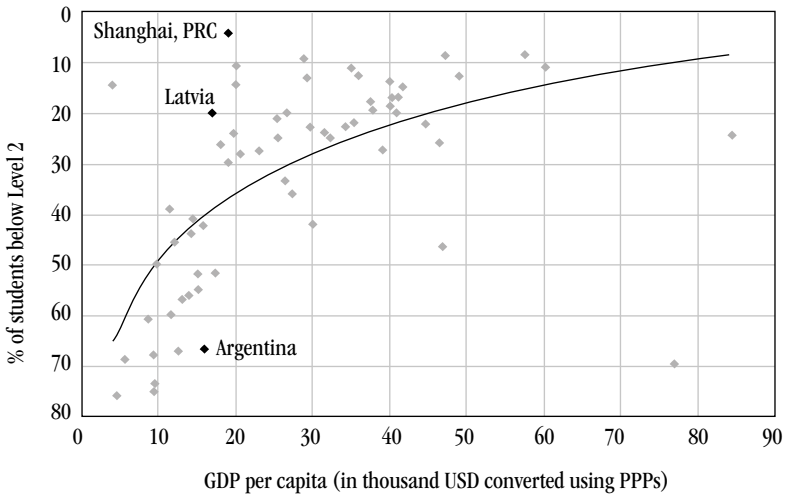


Source: PIAAC database, 2012.

This is confirmed by the data presented in *Figure 12.7*, which compares the number of low-performing students in PISA 2012 with GDP per capita, as a measure of national wealth. While there is a weak overall relationship between national income and the share of low-performing students, there is huge variation among countries, especially among low- and middle-income countries. Compare for example three economies which are situated in more or less the same range of national income: Shanghai (PR China), Latvia, and Argentina. While Argentina counts around 67 per cent of low-performing 15-year-old students in PISA, Latvia has only 20 per cent and Shanghai less than 4 per cent.

It is not a country's national income or financial opportunities that determine whether it has a high or low bottom of the learning pyramid. It is the political will and the determination to include all learners in a process of educational growth and progress that seem to be important.

Figure 12.7 Relationship between national income and share of low-performing students



Source: PISA database, 2012.

Conclusions

In this short paper we explored the bottom of the learning pyramid and its relationship to a number of variables. The first conclusion we can draw is that, with respect to measures of learning, national averages disclose only a very small part of the reality – it is the distribution that really matters. The bottom of the distribution – in either educational attainment or learning outcomes or skills, in absolute terms (the size of the poorly educated or low-skilled population) or in relative terms (the distance from the median or the length of the tail of the distribution) – especially deserves its own metrics.

A second conclusion is that both the share of the low-skilled population and the distance between the bottom and the median of the distribution differ greatly among countries. The latter indicator can be seen as a measure of that country's tolerance for low learning opportunities and outcomes. The level of tolerance for exclusion from learning and education does not seem to be a function of the average level of learning outcomes of a country, or its overall social inequality based on national income, but seems to be an autonomous characteristic of countries.

The third conclusion then, and the more optimistic one, is that the size of the bottom of the learning pyramid and its relative distance from the median are influenced by political decisions and educational policies in particular. There is no general model of educational development, in which 'more of the same' leads to more or less predictable and comparable results. Rather, countries follow their own distinct pattern of educational development, and make political choices that result in very different outcomes, especially with regard to those at the bottom of the learning pyramid. Some countries develop educational policies that really care about those at the bottom of the pyramid and ensure that as few as possible are left behind.

A fourth conclusion is that international intergovernmental organizations like UNESCO and the OECD play a special role with regard to empirical findings, such as those described above. Firstly, they can promote the development and implementation of comparative measures of learning, essential for monitoring the SDG 4 targets. Using PISA, PIAAC, and other international assessments, we have been able to demonstrate how different countries at the bottom are doing. But the second role of international organizations is probably even more important, namely the spread of shared values, the development of a common understanding of educational challenges and, ultimately, the promotion of evidence-based policies through processes of reviewing, peer-learning, and policy advice.

Looking at these data from the perspective of SDG 4, it is clear that a high tolerance for low educational opportunities or learning outcomes is simply not acceptable or sustainable. It will not suffice for a country to bring its average educational attainment or learning outcomes up to desired national standards without at the same time ensuring that the distribution shrinks, especially at the bottom. The risk of policies that promote educational expansion, or international competition on learning outcomes and skills, is that countries will concentrate their efforts on the national average and favour progress at the top. The real challenge of SDG 4 is to promote an inclusive approach to progress in education and learning. The bottom of the pyramid of learning really matters!

References

Benaabdelaali, W.; Hanchane, S.; Kamal, A. 2011. *A new data set of educational inequality in the world, 1950–2010: Gini index of education by age group*. Rabat: Association Marocaine de Sciences Economiques. Accessed 22 June 2017: www.amse.ma/doc/WP_2012_09.pdf

OECD. 2016*a*. *Low-performing students: why they fall behind and how to help them succeed*. Paris: OECD Publishing.

——. 2016*b*. *PISA 2015 Results (Vol. I): excellence and equity in education*. Paris: OECD Publishing.

Van Damme, D. 2014. 'How closely is the distribution of skills related to countries' overall level of social inequality and economic prosperity?' Education Working Paper no. 105. Paris: OECD Publishing. Retrieved from: http://dx.doi.org/10.1787/5jxvd5rk3tnx-en

Commentary

Suzanne Grant Lewis

Van Damme claims that a tolerance for low performance is the result of more or less deliberate policies to leave people behind. While this may indeed be correct, a deeper look at policy design, policy implementation, and resource allocation at the country level is required to draw such a conclusion. What Van Damme calls a policy 'tolerance' for exclusion in learning instead may be the unintended consequence of highly inclusive policies, a disconnect resulting from a number of factors.

An important explanation for poor outcomes from a well-intended policy is the inadequate or inappropriate allocation of resources to implement the policy. Unfortunately, we don't know enough about education finance within countries: where the money comes from, where it goes, or whether it is spent efficiently. IIEP and UIS's work on National Education Accounts with partners in eight countries[38] found some surprising results when educational expenditures were more accurately tracked (UIS and IIEP-UNESCO, 2016).

We found that households, not governments, are the main funders of education in Uganda (57 per cent) and Nepal (49 per cent), and are significant funders in Côte d'Ivoire (33 per cent) and Viet Nam (24 per cent). When families carry so high a burden for their children's education, it is more difficult to ensure equity of access and learning. Since teaching materials are mostly funded by households, children from poorer households are likely to have less access to adequate materials, lowering their chances for higher learning outcomes. Yet, without these details on household expenditure, governments could not consider them in policy design, with policy outcomes more likely to result in large distributions in both per-student expenditure and student learning outcomes. A faulty policy design, not intent, is to blame.

The implementation of central ministry of education policies can also be hindered when local government or government bodies outside the ministry of education provide most of the government's educational funding, as is seen in Viet Nam (84 per cent), Lao PDR (82 per cent), and Uganda (78 per cent). Having more public actors complicates the implementation, and monitoring, of central policy. A

38. The countries are Côte d'Ivoire, Guinea, Lao PDR, Nepal, Senegal, Uganda, Viet Nam, and Zimbabwe.

similar case can be made when external educational financing is not transparent and 'on budget'.

Just as we have insufficient evidence on what interventions in a given context contribute to higher learning outcomes, the lack of accurate data on expenditures by all financing sources means governments are designing inclusive education policies in the dark. Few countries are tracking and analysing financial flows in the education sector; yet by doing so, governments could better design feasible policies to improve the learning outcomes of all children and youth.

Turning from the financial to the political, it is worth considering what additional roles inter-governmental agencies (IGAs), like the OECD and UNESCO, can play to ensure that countries not only bring average learning levels up, but also shrink the distribution of learning outcomes. International bodies governed by members' governments are political by design. They understand that policy-making is fundamentally about the 'competition between multiple social goals and the pursuit of social values' (Parkhurst, 2017: 8). IGAs promote common values and their translation into national and international policy. Since politics also enters into the creation, selection, or interpretation of evidence, as well as the use of research evidence for policy, there is a need for 'the good governance of evidence' as Justin Parkhurst calls it (2017: 159).[39] IGAs are well positioned to play a normative role in supporting both good use of evidence within a policy process and good governance of evidence. This would be a powerful contribution in ensuring learning at the bottom of the pyramid.

References

Parkhurst, J. 2017. *The politics of evidence: from evidence-based policy to the good governance of evidence.* Abingdon, UK and New York: Routledge.

UIS (UNESCO Institute for Statistics); IIEP-UNESCO. 2016. *Who pays for what in education?: the real costs revealed through National Education Accounts.* Montreal, Que.: UIS. Retrieved from: http://uis.unesco.org/sites/default/files/documents/who-pays-for-what-in-education-national-revealed-through-accounts-2016-en_0.pdf

39. This 'requires identifying the set of values dealing with both processes and outcomes that are of particular relevance to the use of evidence within a political process.'

Commentary

Emmanuel Jimenez

Van Damme has three main conclusions: (a) the distribution of attainment or learning in a country is at least as (if not more) important than comparing educational performance; (b) the level of 'tolerance' for exclusion is independent of average learning or overall social inequality in a country; and (c) educational disparity can be influenced by political decisions and policies.

On (a) and (b), it is hard to argue against the Van Damme's proposition that one should focus not only on average attainment and quality in a country, but on distribution as well. His chapter uses the share of those without an upper secondary education (a proxy for educational attainment) and the levels of performance of the lowest-scoring students in PISA (a measure of the adequacy of foundational skills) as a measure of each country's 'tolerance' of letting the bottom lag. Even the relatively well-off OECD countries vary in their performance.

I do have a quibble about the paper's conclusion that the level of tolerance for exclusion is independent of the country's average level of attainment. The paper uses, as a measure of the disparity in educational attainment within OECD countries, the difference between the average annual growth rates – between 2000 and 2013 – of the share of adults without upper secondary education and those with some tertiary education. For some countries, such as Sweden, Finland, Israel, and Canada, this difference is negative and large, so that the gap in adults' educational attainment is narrowing. For others, such as Portugal, Turkey, Italy, and Switzerland, the difference is positive and large. The paper implies that the former countries prioritized the needs of the bottom relative to the latter.

While interesting, it is difficult to interpret any policy intent or effort to prioritize investments from these numbers. For example, it is easier for a country at a lower initial level to increase educational attainment relative to others. Consider two of the countries that are considered to have widened educational disparity (*Figure 12.2*). Turkey increased its share of tertiary by 5 per cent and Switzerland by just above 4 per cent; both decreased their share of those without upper secondary by, respectively, 1.5 and 2 per cent. So, by the paper's metric, Turkey's prioritization led to more disparity (5 per cent –1.5 per cent = 3.5 per cent) relative to Switzerland's (4 per cent – 2 per cent = 2 per cent). But Turkey's initial level of tertiary

enrolment, from the OECD Education Database, was only 18 per cent, while Switzerland's was already a whopping 42 per cent. Shouldn't Turkey be given more credit than Switzerland for lowering disparity?

On (c), I also have to agree. But I wish the paper could have been longer, because there is very little discussion on how to address politics and policies. What is a country to do? What may come immediately to mind is more money. The Education Commission (2016) estimated that, to meet the UN goals by 2030, education spending needs to increase threefold from the $1.2 trillion being spent now; but that report also admitted that money isn't everything. Governments spend '2 percent of GDP on education costs that do not lead to learning' (Education Commission, 2016: 7). And what they do spend is not prioritized on the bottom of the pyramid.

Van Damme rightfully calls for the 'promotion of evidence-based policies'. The main challenge is that the number of studies that are able to attribute outcome to an intervention remains pitifully small. The most recent of several systematic reviews on education done by 3ie, which included randomized controlled trials and other rigorous methods that address causality, found only 238 studies of 214 interventions from 1990 to 2015 across all low and middle income countries (Snistviet *et al.,* 2016). Many of these studies were focused on particular interventions, such as cash transfers. Other interventions that can affect the learning of the BoP, and which consume large amounts of resources, are under-studied.

Moreover, 3ie's report, as well as other recent systematic reviews, have found no magic bullets among the interventions that have been studied (Evans and Popova, 2015). 'Structured pedagogy' interventions, which invest in a package of curricula, pedagogic techniques, appropriate materials, and teacher training that are tailored to local barriers to education, do lead to significant enhanced learning. But most other interventions have low effect sizes because of the large variances across studies – local context matters. This means that policy-makers need to build evaluations into their local interventions and learn by doing.

More in-depth evaluations of what works to address the needs of the bottom billion are needed. They should focus not only on the effects of educational interventions, but also on the sector's governance of the sector because, as Van Damme states, learning for these people is a political issue. Without insights gained from such research, the answer to the chapter's title may be that it is, unfortunately, sustainable to leave the bottom behind.

References

Education Commission (International Commission on Financing Global Educational Opportunity). 2016. *The learning generation: investing in education for a learning world.* New York: Education Commission. Retrieved from: http://report.educationcommission.org/report/

Evans, D.K.; Popova, A. 2015. 'What really works to improve learning in developing countries: an analysis of divergent findings in systematic reviews', Policy Research working paper no. WPS 7203. Washington DC: World Bank. Retrieved from: https://openknowledge.worldbank.org/handle/10986/21642

Snistviet, B.; Stevenson, J.; Menon, R.; Philips, D.; Gallagher, E.; Geleen, M.; Jobse, H.; Schmidt, T.; Jimenez, E. 2016. *The impact of education programmes on student participation and learning in low and middle income countries, 3ie systematic review summary number* 7. London: International Initiative for Impact Evaluation (3ie). Retrieved from: www.3ieimpact.org/en/publications/systematic-review-publications/3ie-systematic-review-education-effectivenes-srs7/

The International Institute for Educational Planning

The International Institute for Educational Planning (IIEP) is an international centre for advanced training and research in the field of educational planning. It was established by UNESCO in 1963 and is financed by UNESCO and by voluntary contributions from Member States. In recent years the following Member States have provided voluntary contributions to the Institute: Argentina, France, Norway, Sweden, and Switzerland.

The Institute's aim is to contribute to the development of education throughout the world, by expanding both knowledge and the supply of competent professionals in the field of educational planning. In this endeavour the Institute cooperates with training and research organizations in Member States. The IIEP Governing Board, which approves the Institute's programme and budget, consists of a maximum of eight elected members and four members designated by the United Nations Organization and certain of its specialized agencies and institutes.

Inquiries about the Institute should be addressed to:
The Office of the Director, International Institute for Educational Planning,
7–9 rue Eugène Delacroix, 75116 Paris, France