



Improving Quality Education and Children's Learning Outcomes

and Effective Practices in the Eastern and Southern Africa Region
Report for UNICEF ESARO

MAIN REPORT



Australian Council for Educational Research



Tim Friedman, Ursula Schwantner, Jeaniene Spink, Naoko Tabata and Charlotte Waters
Australian Council for Educational Research (ACER)
2016

Commissioned by UNICEF Eastern and Southern Africa Regional Office (ESARO), Basic Education and Gender Equality (BEGE) Section.

©2016 United Nations Children's Fund (UNICEF)

Cover photo: Françoise d'Elbee ©UNICEF Kenya/2009/d'Elbee

Permission is required to reproduce any part of this publication. Permission will be freely granted to educational or non-profit organizations. Others may be requested to pay a small fee. Requests should be addressed to: Basic Education and Gender Equality Section, Eastern and Southern Africa Regional Office, UNICEF; Tel: + 254 207-622-307 email: jolang@unicef.org.

Contents

Lists of tables and figures	ii
Acronyms and abbreviations	iv
Acknowledgements	vii
Foreword from the Eastern and Southern Africa Regional Director	viii
Executive summary	x
Introduction	1
Context for primary education in the ESA region	1
Conceptual framework	2
1. Stock-taking and comparative analysis of existing assessments in the ESA region	5
1.1 Overview of assessments in the stock-taking	5
1.2 Comparative analysis of assessments	11
2. Literacy and numeracy in primary education in the ESA region: Students experiencing limited learning outcomes and trends over time	27
2.1. Characteristics of students experiencing limited learning outcomes in literacy and numeracy in primary education in the ESA region	28
2.2. Trends in learning outcomes of children in primary education in literacy and numeracy in the ESA region	38
3. Improving learning outcomes in the ESA region: Effective country-level practices	45
3.1 Country-level programmes analysed	46
3.2 Key strategies for success	47
4. A macro theory of change	51
References	57

Lists of tables and figures



Tables

Table 1: Participation in international assessments in ESAR countries	6
Table 2: Participation in regional assessments in ESAR countries.....	7
Table 3: Implementation of national assessments by government/parastatal bodies in ESAR countries.....	8
Table 4: Implementation of EGRA/EGMA in ESAR countries.....	10
Table 5: Purpose of the assessments from stock-taking	11
Table 6: Target population in the assessments from stock-taking.....	12
Table 7: Cognitive domains in assessments from the stock-taking	15
Table 8: Test administration methods of assessments from the stock-taking	18
Table 9: National assessments and use of IRT in data analysis	20
Table 10: Descriptions and attainment of competency levels in literacy: NASMLA in Kenya 2010	21
Table 11: Proportions of students defined as experiencing limited learning outcomes by assessment, country, grade and domain	30
Table 12: Direction of trends in SACMEQ reading and mathematics scale scores from 2000 to 2007, based on (Hungu et al., 2011)	43

Figures

Figure 1: A macro theory of change. An evidence-based monitoring and intervention cycle as premise for change: assessment, analysis, action	xiv
Figure 2: Assessments from the stock-taking by type of assessment	6
Figure 3: Mean percentage marks in Grade 3 mathematics by gender and province – Annual National Assessment in Kenya	22
Figure 4: Trends in English performance across time for students in Kenya (Uwezo)	39
Figure 5: Trends in mathematics performance across time for students in Tanzania (Uwezo) .	40
Figure 6: Trends in proportions of students experiencing limited learning outcomes across Uwezo countries.....	41
Figure 7: A macro theory of change. An evidence-based monitoring and intervention cycle as premise for change: assessment, analysis, action	52

Acronyms and abbreviations



3ie	International Initiative for Impact Evaluation
ACER	Australian Council for Educational Research
ANA	Annual National Assessment
ASALs	Arid and semi-arid lands
ASLI	Africa Student Learning Index
BEGE	Basic Education and Gender Equality
CBA	Competency-based approach
CO	Country Office
CONFEMEN	Conference of Ministers of Education of French-speaking Countries
DAC	Development Assistance Committee
DfID	Department for International Development (UK)
ECD	Early childhood development
EDF	Education Development Fund
EDI	Education for All Development Index
EFA	Education for All
EFA GMR	Education for All Global Monitoring Report
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
ELMI	Early Literacy and Maths Initiative
EMIS	Education Management Information System
ESA	Eastern and Southern Africa
ESAR	Eastern and Southern Africa region
ESSP	Education Sector Strategic Plan
ETF	Education Transition Fund
GPE	Global Partnership for Education
GPI	Gender Parity Index
IEA	International Association for the Evaluation of Educational Achievement
IEP	Integrated Education Programme
IfE	Innovation for Education
IIEP	International Institute for Educational Planning
IRT	Item response theory
LARS	Learning Achievement in Rwandan Schools
LLO	Limited learning outcomes
LNAEP	Lesotho National Assessment of Educational Progress
MICS	Multiple indicator cluster surveys
MLA	Monitoring learning achievement

MoESAC	Ministry of Education, Sport, Arts and Culture of Zimbabwe
MTPDS	Malawi Teacher Professional Development Support
NALA	National Assessment of Learner Achievement
NAPE	National Assessment of Progress in Education
NASMLA	National Assessment System for Monitoring Learning Achievement
NER	Net enrolment rate
NLA	National Learning Assessment
NSAT	National Standardized Achievement Test
OECD	Organisation for Economic Co-operation and Development
OVC	Orphans and vulnerable children
PASEC	Programme for the Analysis of the Education Systems of CONFEMEN Countries
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PRIMR	Primary Math and Reading Initiative
RCT	Randomised control trial
READ	Russia Education Aid for Development
REB	Rwanda Education Board
RTI	Research Triangle Institute
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SBA	School-based assessment
SSME	Snapshot of School Management Effectiveness
TAC	Teacher Advisory Centre
TIMSS	Trends in International Mathematics and Science Study
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
UNICEF ESARO	UNICEF Eastern and Southern Africa Regional Office
USAID	United States Agency for International Development
ZELA	Zimbabwe Early Learning Assessment
ZIMSEC	Zimbabwe School Examinations Council



Acknowledgements

The Australian Council for Educational Research (ACER) was contracted by the United Nations Children’s Fund (UNICEF) to deliver a consultancy service for improving the quality of education and children’s learning outcomes and effective practices in the Eastern and Southern Africa region. This study was co-financed by UNICEF and ACER’s Centre for Global Education Monitoring.¹ The authors of this report would like to thank Manuel Cardoso, Education Specialist, Programme Division, UNICEF; Camille R. Baudot, Regional Adviser, Basic Education and Gender Equality (BEGE), UNICEF Eastern and Southern Africa Regional Office (ESARO); Mitsue Uemura, Shiraz Chakera, Inge Vervloesem, Benoit d’Ansembourg and Pablo Stansbery, Education Specialists from BEGE, UNICEF ESARO, as well as the focal persons at the 21 UNICEF Country Offices in the Eastern and Southern Africa Region (ESAR) for their support during the conduct of the study and preparation of the report.²

We would also like to thank the key informants who took part in the interviews conducted as part of this study: John Mugo (Regional Director of Uwezo at Twaweza); Ben Piper (Lead in the Primary Maths and Reading Initiative, PRIMR, in Kenya) and the Kenya PRIMR team; Kenneth Russell (Education Specialist, UNICEF Zimbabwe Country Office); Professor Robert McCormick (Innovation for Education Evaluation and Monitoring advisor); and Joyce Kinyanjui (Programme Manager, Women Educational Researchers of Kenya, Opportunity Schools).³

Thanks also to the many other individuals who provided information for this study, including Aaron Benavot (Director, GMR team, UNESCO) and Nihan Blanchy-Koseleci (UNESCO GMR); Scott Murray (DataAngel); Marc van der Stouwe (Mott MacDonald/Cambridge Education); Lucy Maina (Regional Director of Africa Educational Trust, Somalia); and Vyjayanthi Sankar (International Consultant, Quality Education, UNICEF ESARO). We would also like to acknowledge the contributions of several education authorities in the ESA region, including the Botswana Ministry of Education and Skills Development; the Botswana Examinations Council; the Ethiopia National Educational Assessment and Examinations Agency; the Examinations Council of Lesotho; the Madagascar Ministry of Education; the Malawi Ministry of Education, Science and Technology; the Mozambique Ministry of Education; the Directorate of National Examinations and Assessment Namibia; the Director-General Department of Basic Education of South Africa; the Uganda National Examinations Board; the Examination Council of Zambia; and the Zimbabwe School Examinations Council.

The authors of this report are Tim Friedman, Ursula Schwantner, Jeaniene Spink, Naoko Tabata and Charlotte Waters, with valuable contributions from Elizabeth Cassity, Mary Kimani, Alejandra Osses, and Adeola Capel.

1 For further information about ACER’s Centre for Global Education Monitoring (GEM) visit <https://www.acer.edu.au/gem>.

2 The Eastern and Southern Africa (ESA) region, as under UNICEF programming, encompasses 21 countries: Angola, Botswana, Burundi, the Comoros, Eritrea, Ethiopia, Kenya, Lesotho, Madagascar, Malawi, Mozambique, Namibia, Rwanda, Somalia, South Africa, South Sudan, Swaziland, the United Republic of Tanzania (later in this report referred to as ‘Tanzania’), Uganda, Zambia and Zimbabwe.

3 Not all the interviews conducted for this study were included in the report. However, the contributions were highly valuable and the authors thank all interview partners for their time and information.

Foreword from the Eastern and Southern Africa Regional Director

Over the past few decades the world's attention has been focussed on attaining Millennium Development Goal 2 – universal access to primary education. During this period, governments and the international community have been investing on school infrastructure, training teachers, and learning materials. For UNICEF globally and across Eastern and Southern Africa, the challenge of our time is now how to sustain the momentum in access and to reinforce quality learning outcomes. There is, today, an urgent global realisation that beyond getting children in the classroom, it is imperative that they learn.

The new Sustainable Development Goal (SDG) 4, presents huge opportunities to meet this challenge through a strategic shift towards equitable quality education for all. This shift is essential. Emerging evidence shows that large numbers of children are in school, but are not learning. In 2012, the Africa Barometer report by the Centre for Universal Education at Brookings estimated that of the 97 million who enter school on time in Sub-Saharan Africa, 37 million will not learn basic skills. Thirty seven million: that is one-third of all children who go to school will reach their adolescent years unable to read, write, and/or perform basic numeracy tasks.

The analysis from *Improving quality education and children's learning outcomes and effective practices in the Eastern and Southern Africa region* has similar conclusions. It reveals that as many as 40 per cent of children in school do not reach the expected basic learning benchmarks in numeracy and literacy. The new report also confirms that children from families with lower socio-economic status and whose home language is different from the language of instruction are less likely to learn.

The findings – that many children are in school, but not learning – represent a huge waste of human and financial resources. Fundamentally, the promise of education and the transformative opportunity of schooling for children, families and communities is not being fulfilled.

UNICEF believes that confronting this learning crisis, through high impact solutions, is the priority for the education community. Encouragingly, the report highlights that many countries in East and Southern Africa are promoting quality through improving learning monitoring by national, regional and international learning assessments, and through developing targeted programmes that improve teaching and learning.



Given this learning crisis, acceleration of these trends is essential. UNICEF will push for an increased system-wide emphasis on outcomes as opposed to inputs; improved assessments to gauge children's learning progress; and building knowledge of the pedagogic practices that can improve learning.

There is so much potential. While highlighting challenges, this report also shows elements of national progress. UNICEF encourages countries to accelerate these developments. We are cognisant of the many challenges facing education in the region – 1 in 5 children not attending school; a demographic boom in the region that will see 70 million additional children by 2030; and continued overstretched public finances. In such critical circumstances, the winning combination of access, quality learning, and affordability is ever more crucial.

It is in this context that the report provides us with a critical baseline on quality education in every country in the region. The report assesses the learning outcomes being reached in the region, the learning assessment tools countries are deploying to generate evidence on learning, and the interventions countries are implementing to improve teaching and learning.

With this report, UNICEF and our many partners will be better equipped to support improvements in quality education for children.

A handwritten signature in blue ink, appearing to read 'Leila Gharagozloo-Pakkala', with a long, sweeping underline that extends downwards and to the right.

Leila Gharagozloo-Pakkala
Regional Director
Eastern and Southern Africa
United Nations Children's Fund

Executive summary



Introduction

The Eastern and Southern Africa (ESA) region is progressing well towards achieving important Education for All (EFA) goals, particularly with regard to increasing student enrolment in the primary years (UNESCO, 2014). Despite this achievement, there is still considerable work to be done to improve the quality of education. Primary school students in low-income sub-Saharan African countries have, on average, learned less than half of what is expected of them (Majgaard and Mingat, 2012, p.6). The gap between the learning achievements in developed economies and the learning achievements in Eastern and Southern Africa is estimated to be at least four grades (GPE, 2012, p.116).

In order to understand the major impediments to student learning in the region, the United Nations Children's Fund (UNICEF) contracted the Australian Council for Educational Research (ACER) to take stock of and compare existing student assessments in the region, focusing on students in primary education. The terms of the contract called for ACER to study the existing assessment systems and methodologies in the region, and document how the assessment data are derived and used to inform education policy in the region. We were also asked to identify factors and practices that could help improve learning outcomes in literacy and numeracy in primary education, specifically for disadvantaged children with limited learning outcomes (LLOs).

Our study consists of three research components. The first provides an overview and comparative analysis of the existing assessments of student learning outcomes in literacy and numeracy in primary education in the region. The second considers the characteristics of children experiencing LLOs in the domains of literacy and numeracy, including trends in achievement over time. The third looks at effective country-level practices in the ESA region that could improve learning outcomes in the literacy and numeracy of disadvantaged children in primary education. Our report concludes with a macro theory of change drawing on the evidence we gathered for this report.

Comparative analysis of existing assessments in the ESA region

Our study covered 23 countries and identified 58 existing assessment systems that evaluate student learning outcomes in literacy and numeracy in primary education. Of these, EGRA and EGMA are the most prevalent programmes (36 per cent) followed by regional (29 per cent), national (28 per cent) and international (7 per cent) assessments. Most of these assessments target lower-primary students (grade 2 or 3) and most commonly focus on literacy and numeracy assessments. While these commonly use mean scores for the cognitive results or frequency analysis for the contextual data, item response theory (IRT) methods, which can scale data and meaningfully compare results across grades, contexts and time, is less prevalent. Contextual data linked to the cognitive results is available for many, but not all of the assessments, making it difficult to draw policy-related findings from the results. Not captured in our stock-taking was whether, and to what extent, student assessment data is linked to Education Management Information System (EMIS) data on a systems level. Access to the data is a challenge, and our study found that while the results of 71 per cent of the assessments were published, we were unable to obtain the original datasets for the remainder.

Students experiencing LLOs and trends over time

The objective of the study was to investigate the characteristics of children experiencing LLOs and trends in their performance over time. In 32 out of the 58 assessments, competency-level benchmarks are defined. However, each of the assessments (PASEC, UWEZO, TIMSS and prePIRLS) used different metrics for literacy and numeracy. Therefore, there is no shared benchmark among them that could be used to construct a common definition of 'limited learning outcomes.' Instead, we employed the benchmarks each assessment used to gauge literacy and numeracy. Given the differences in these metrics across data sets, countries and year level, the percentage of students identified with LLOs is wide ranging, from 18 per cent to 40 per cent in numeracy, and 18 per cent to 50 per cent for literacy.

In international and regional assessments for the ESA region, average test scores for literacy and numeracy are generally low, with a considerable percentage of students failing to have acquired basic skills in reading and mathematics. In Lesotho, for example, by Grade 6 only 48 per cent of students have achieved basic reading skills. In Zambia and Malawi, only 27 per cent of students achieved this level. In mathematics, the proportion of primary students with basic skills is considerably lower, with fewer than 50 per cent of students in Grade 6 achieving the minimum level in two-thirds of the countries (UNESCO, 2014, p.35).

Consistent with other studies conducted in the region, individual and family characteristics of students, such as gender, age, language spoken at home, socio-economic factors, preschool attendance, activities prior to attending school, engagement and out-of-school tuition, were all found to be associated with the likelihood that a student would experience LLOs in literacy or numeracy. In addition, the type of school, the location of the school and the resourcing available to the school that the student attends also contributes to the likelihood that the student would be experiencing LLOs.

Our study showed that, in general, males are more likely to experience LLOs in literacy than females. In Botswana, for example, males are almost three times more likely than females to experience LLOs. While on the whole girls outperformed boys in reading literacy, rural boys outperformed rural girls on almost all tasks, but in urban schools the opposite was the case (RTI, 2010, p.37). However, in mathematics, on the whole, boys outperformed girls.

The age of the student relative to the school entry grade is an equally important factor. While the relations between age and performance are complex and may be determined by different socio-economic and demographic factors, Hungi et al (2014) found that in developed countries older students generally outperform their younger colleagues, while in developing countries, especially in Africa, younger students perform better than older students. Our study supported this insight. Across the region, we found that students who were relatively younger than the median class age tended to be less likely to be experiencing LLOs. For instance, Grade 6 students in Botswana who were 12 years or less were almost three times less likely to be experiencing LLOs in mathematics than students 12 years of age or older.

The language spoken at home also has a strong impact on learning outcomes. In countries where the official language is not the most common language spoken at home, there are strong links between language and marginalisation in education. Evidence from PASEC and SACMEQ show a strong link between home language and the language of instruction in determining test scores (Fehrler and Michaelowa, 2009, UNESCO 2010, p. 154; Garrouste, 2011). While low language skills are commonly viewed as a critical factor in literacy assessments, evidence from Namibia using SACMEQ results suggests that they also make a significant contribution to low performance levels in mathematics (Garrouste, 2011, p. 231).

Furthermore, the socio-economic status of students is a strong predictor of achievement. Our study found that students from lower socio-economic backgrounds were more likely to experience LLOs across all countries examined in both literacy and numeracy. We found this to be the case across all countries, despite the different measures used to assess socio-economic status. Among other factors, household possessions, including the availability of reading materials and books in the home, and levels of parental education, were also found to be associated with LLOs. Furthermore, the amount of time that students spent working was negatively associated with achievement data (ACER and ZIMSEC, 2015).

Students who had limited exposure to a learning environment in the home were disadvantaged in performance at school. A profound impact on learning outcomes was evident in homes where students were involved in reading and storytelling, were not required to work outside of the home, started school early and were provided with adequate support in school by their teachers to build foundational literacy skills, and attended schools that had relevant and engaging reading and learning materials in buildings with clean water and sanitation.

Our study also considered student performance trends over time. While much of Eastern and Southern Africa has experienced a marked improvement in student enrolment, student performance has changed little over time. Indeed, on the whole, student performance has stagnated or worsened. It must be noted, however, that given the limited comparable data available for our study, drawing any general conclusions for the whole ESA region is problematic. Instead, conclusions based on improvement or decline in student abilities should only be considered at the national level.

Effective country-level practices

As part of our study, a number of strategies were identified that contributed to the success of country-level practices in the ESA region. We found that while there is a considerable body of literature for the ESA region on practices that increase quantitative aspects of education quality, such as access, enrolment and retention rates, we found few reports on programmes to improve student learning outcomes.

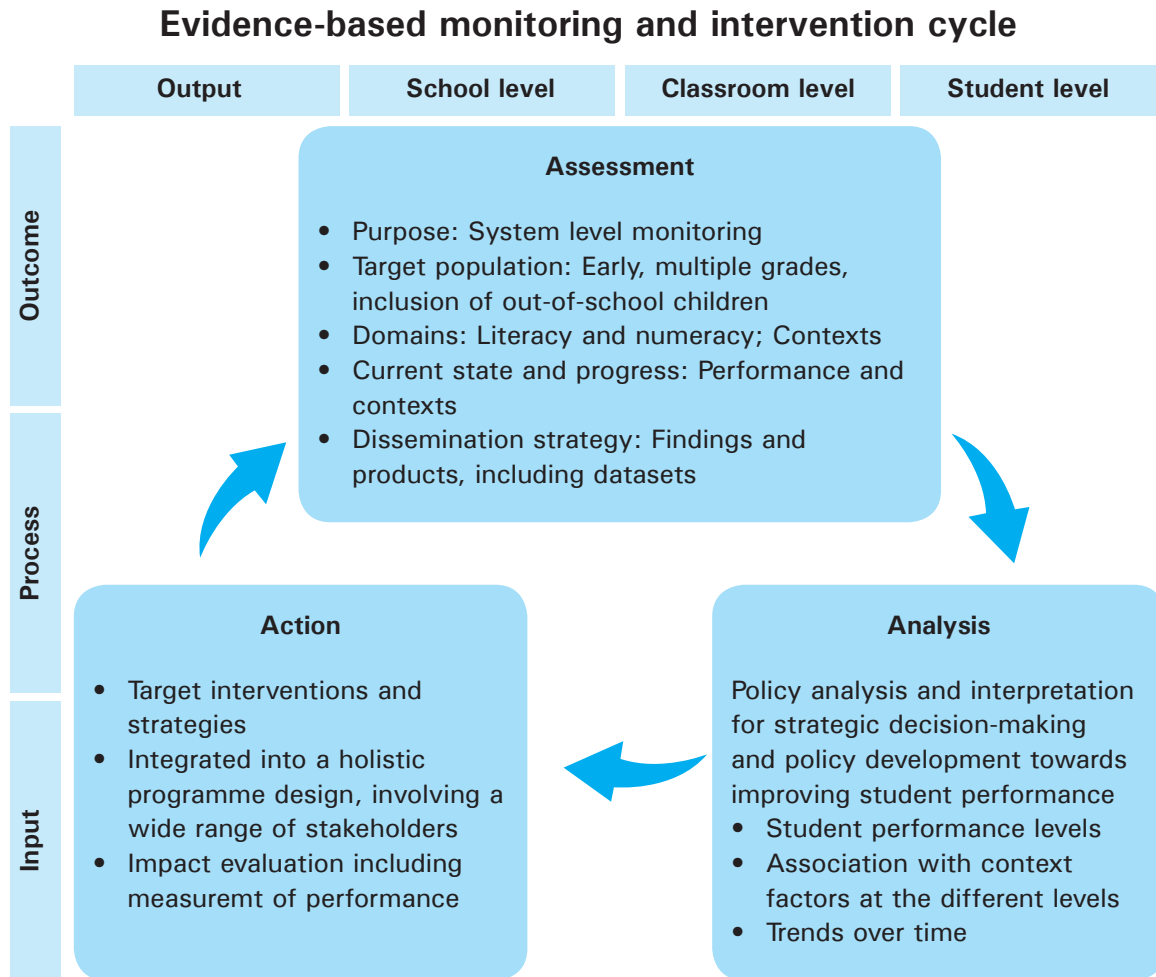
Altogether, we identified 10 programmes in 7 out of the 21 countries identified as having had an impact on student learning in early grade literacy/numeracy. These comprised of teacher training on reading/mathematics instruction; provision of teaching-learning materials; production of reading materials in the local language; and community- and home-based reading activities which were linked to effective ECD programmes. Additionally, programmes that aimed at a whole-school improvement strategy were shown to have a significant impact on learning outcomes.

Broadly, these successful programmes use a three-pronged approach comprised of assessments, teacher training and community support for children's reading. They provide a combination of well-targeted instructional interventions, regular professional development of teachers through school-level training and coaching, with regular system-level follow up and support, matched with sufficient relevant and quality classroom materials, and more literacy and numeracy instructional time. Students having a reading buddy to support their learning to read had a positive effect on learning outcomes in a number of locations. Overall, our study found that key strategies for improving learning outcomes of disadvantaged children share two common features: a holistic and coherent approach and consistent and continuous support over time.

A macro theory of change

Based on the evidence collected for our study, we developed a macro theory of change aiming at monitoring and improving literacy and numeracy performance of children in primary education in the region. The theory combines the main findings of each stage of the study and highlights the '3As' approach for long-term and sustainable change in student performance: assessment, analysis and action. Critical to this framework is the dissemination of the assessment results in order to initiate action by governments, communities, parents and development partners (see Figure 1).

Figure 1. A macro theory of change. An evidence-based monitoring and intervention cycle as premise for change: assessment, analysis, action



Conclusions

Our study analysed existing student assessments, data resulting from some of these assessments and effective country-level programmes in the ESA region.

We found that programmes targeted toward early learning in disadvantaged communities made the biggest impact. The level of exposure students have to a learning environment, either through home or school, in their early years and the presence of holistic, system-level educational programmes that support quality early learning programmes in disadvantaged communities made a significant contribution to improved student performance.

While many researchers have studied the factors that contribute to student school attendance, fewer have explored what helps improve student learning. In financially constrained environments, resources should be targeted at understanding the gaps in the system with regard to student performance and supporting effective interventions.

In order to do this, policymakers must consider how learning assessment programmes that provide quality comparable data across population subsets, between grades and over time, can be integrated from the outset into education reform agendas.

Introduction

Context for primary education in the ESA region

The Eastern and Southern Africa (ESA) region provides many challenges for primary education. The main issues are poverty, health and social issues, and fragile political/economic circumstances—particularly in Burundi, the Comoros, Eritrea, Madagascar, Somalia and South Sudan (UNESCO, 2014, p. 12).^{4,5}

Overall, the region is making progress toward achieving Education for All (EFA) goals (UNESCO, 2014). The EFA Development Index shows a considerable improvement between 2000 and 2012, with increases in the primary education completion rate, the literacy rate for those 15 years and older, and the primary enrolment of girls and boys (UNESCO, 2014, p. 16; 2015a).

However, very little attention has been paid to the quality of education, or to progress in student performance. While improvements in education can be assessed by quantitative aspects, such as access to education, enrolment and completion rates, or gender parity, such metrics don't assess education quality or, even more importantly, the improvement of student performance.

In most sub-Saharan African countries, average test scores in international/regional assessments of student learning are low. Primary school students in low-income, sub-Saharan African countries have, on average, learned less than half of what is expected of them (Majgaard and Mingat, 2012, p. 6). A comparison of high- and low-income countries using data from the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and the Progress in International Reading Literacy Study (PIRLS) reveals large differences between the poorer economies in SACMEQ (for example, Lesotho, Malawi and Zambia) and the mainly high-income economies in PIRLS.^{6,7}

The gap between the learning achievements in developed economies and the learning achievements in East and Southern Africa is estimated to be at least four grades (GPE, 2012, p. 116).⁸

4 Most countries in the region (16 out of 21) are classified as least developed countries under the United Nations (UN) definition <<http://unohrls.org/about-ldcs/>>. Kenya is among low income countries; Swaziland forms part of the lower-middle income country group; and Botswana, Namibia and South Africa are among upper-middle income countries (according to DAC-ODA-recipient status (OECD Development Assistance Committee, Official Development Assistance). Source: DAC. List of ODA Recipients effective as at 1 January 2015 for reporting on 2014, 2015 and 2016 flows available at <<http://www.oecd.org/dac/stats/daclist.htm>>. However, there are also major disparities within these middle income countries; see <http://www.unicef.org/esaro/theregion_old.html>.

5 <<http://www.worldbank.org/en/topic/fragilityconflictviolence>> viewed 5 May 2015; in 2014, Malawi was also categorised as a fragile state.

6 ESAR countries participating in SACMEQ are Botswana, Kenya, Lesotho, Malawi, Mozambique, Namibia, South Africa, Swaziland, Tanzania, Uganda, Zambia and Zimbabwe.

7 Ross (2009) scaled PIRLS and SACMEQ data using common items and a Rasch model to put the test results on the same scale, based on anchor items and test equating, thus making the data comparable across economies (GPE, 2012, p. 116).

8 Considering that PIRLS measures learning outcomes of primary school students in Grade 4, and SACMEQ in Grade 6, this potentially corresponds to two more grades of schooling (GPE, 2012, p. 116).

For this study we reviewed the assessment systems in the region that measure learning outcomes in literacy and numeracy of children in primary education, and undertook a comparative analysis of these assessments. We analysed existing data from the region to develop a portrait of students who are experiencing LLOs in literacy and numeracy, and examined trends in performance over time. We investigated education interventions that helped to improve students' learning outcomes in literacy and numeracy, especially for disadvantaged children, to identify strategies to further improve student performance in the region. Using these elements, we developed a theory of change that sets out processes for improved learning outcomes.

Conceptual framework

Policies aiming at monitoring and improving educational progress must be based on data on learning outcomes and the factors related to the outcomes. 'Learning outcomes', as defined in this study, refer to student performance in a cognitive domain, in particular literacy and numeracy, as measured in the different assessments discussed in Chapter 1. The conceptual framework we used sets out the contextual factors associated with student performance, which provides a basis for the definition of factors to be considered in an assessment and support for the decision on which level of the education system that innovations are most needed.

The conceptual framework combines the main characteristics of two models that have both been highly influential in the field: the 'Input-Process-Outcome model' (Purves, 1987) and the 'dynamic' model of educational effectiveness (Creemers and Kyriakides, 2008; Kyriakides and Creemers, 2006).^{9,10} In these models, input, process and outcome factors operate at the different levels of an education system, i.e., system-level (national, regional and community); as well as school-, classroom- and student level. Input factors mainly refer to structural conditions, for example, economic wealth or community infrastructure (system level); school type (public, private); school location (rural, urban); and school resources (school level); class-size and teaching resources (classroom level); as well as individual (e.g. gender, age, grade) and family factors, such as socio-economic status and parental education (student level). Process factors mainly concern policies and strategies, and range from national curriculum and teacher education (system level), through management and leadership (school level), quality of instruction (classroom level), to the actual learning process (student level). Outcome factors—i.e., performance in literacy and numeracy—are measured at the student level and can be aggregated at the system level (national, regional and community level), school level, or classroom level.¹¹ Relations between the factors at the different levels are complex and have not yet been fully investigated. In the 'dynamic model,' educational outcomes can become inputs for further development. For example, domain-related attitudes and beliefs can be considered as outcomes of schooling or as inputs affecting student behaviour (Creemers and Kyriakides, 2008; Kyriakides and Creemers, 2006; OECD, 2013).

9 The basic structure of the Input-Process-Outcome model was developed in the 1960s for the International Association for the Evaluation of Educational Achievement (IEA) (Purves, 1987).

10 Other recent examples where the main characteristics of these two models are combined are the contextual framework for OECD Programme for International Student Assessment (PISA) (OECD, 2013), or the input-process-output model for data in the student learning environment (Biggs, 1999; Biggs and Moore, 1993).

11 Other outcome factors at the system, school, and classroom level relate to aggregated pass rates and graduation rates, enrolment and retention rates; at the student level, domain-related and general school-related attitudes, beliefs and motivation are considered important outcomes of schooling.

From an assessment-design perspective – regarding data collection of input, process and outcome factors on the four levels – it may not always be possible, nor feasible, to represent the full framework. The main source for information on input and process variables at the student and school level are questionnaires, sometimes combined with qualitative methods, such as classroom or school observation. Data on input and process factors on the system level are collected at the system level, using tools such as the Education Management Information System (EMIS). Outcome factors, i.e., performance in literacy and numeracy, are measured at student level using cognitive tests.

The comprehensive conceptual framework allows the main effectiveness factors to be considered in an assessment. The framework also has potential to inform the development of theories of change: Input, process and outcome factors can be defined and analysed for their stability and suitability for change. The framework can also inform decisions about the level at which innovations need to be implemented to ensure their maximum effectiveness. For example, determining which input, process and outcome factors need to be addressed by policy decisions at the system and school level, and which can be best shaped by school development activities at the school and classroom level.



1. Stock-taking and comparative analysis of existing assessments in the ESA region

A wide range of assessment systems is in use in the ESA region.¹² For a better understanding of these assessments, the available data and how they support policies aiming at monitoring and improving learning outcomes, we started with an overview of the systems. The focus was on assessments that measure student performance in literacy and numeracy in primary education. Following the overview, we analysed and compared the assessments, highlighting the strengths and limitations of the different methodological approaches. Case studies for Zimbabwe and Rwanda offer an in-depth understanding of specific measurement practices implemented by the governments of these countries (see Appendix II). The detailed results of the stock-taking are presented in the main stock-taking table (see Appendix VI).

1.1 Overview of assessments in the stock-taking

The stock-taking considered assessments that were implemented from 2007 up to 2014/2015 in the ESA region.¹³ Overall, we identified 58 assessments that assess student performance in literacy and numeracy in primary education in the ESA region.

The assessments can be grouped into four types:

- international assessments
- regional assessments
- national assessments
- Early Grade Reading Assessment (EGRA)/Early Grade Mathematics Assessment (EGMA).

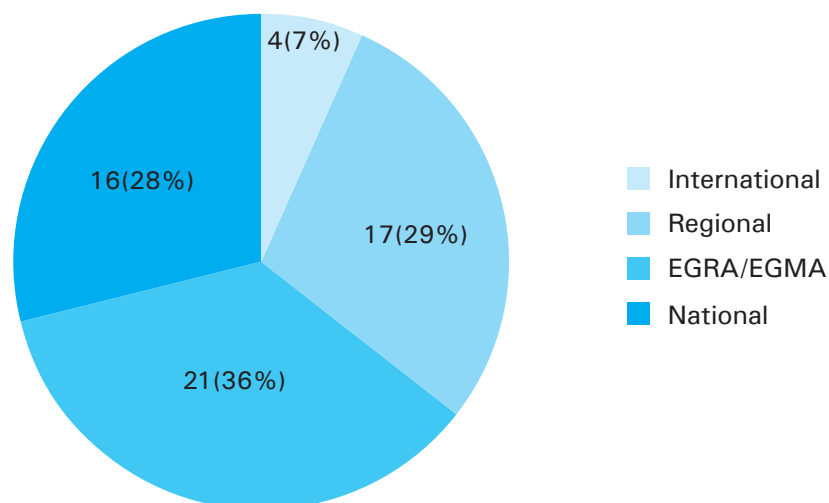
Of the 58 assessments, 4 (7 per cent) are international and 17 (29 per cent) are regional. Implementation of EGRA and EGMA accounts for 21 (36 per cent), which is the largest group among the four types, and 16 (28 per cent) are national (see Figure 2).¹⁴

12 National examinations are not included in this study. Examinations don't share the same purpose as learning assessments, which leads to different choices about sampling, data analysis, reporting (e.g. on examination pass rates and/or achieved grades, i.e. pass levels) etc.

13 The stock-taking of assessments took place between October 2014 and March 2015 and considers assessments that have been implemented since 2007. Assessments conducted outside this period were not considered. In the case of reoccurring assessments, the inception date of an assessment can be before 2007.

14 All percentages are rounded.

Figure 2. Assessments from the stock-taking by type of assessment



Overall, 20 countries have implemented one or more assessments (see Table 16 in Appendix II). Six countries have used several types of assessments (i.e., Kenya, Malawi, Mozambique, South Africa, Uganda and Zambia), while another six countries show limited assessment activities with one or no assessment implementation in recent years (i.e., Angola, Comoros, Eritrea, Madagascar, South Sudan and Swaziland) (see Table 17 in Appendix II).¹⁵

Participation in international assessments

Two countries in the region have participated in international assessments of the IEA (International Association for the Evaluation of Educational Achievement): TIMSS (Trends in International Mathematics and Science Study), PIRLS (Progress in International Reading Literacy Study) and prePIRLS (see Table 1).¹⁶

Table 1. Participation in international assessments in ESAR countries

Country	Assessment
Botswana	TIMSS in 2007, 2011 PIRLS in 2011, prePIRLS in 2011
South Africa	TIMSS in 2011 PIRLS in 2006, 2011; prePIRLS in 2011

Since 1995, TIMSS has been measuring trends in mathematics and science achievement at Grade 4 and Grade 8. Where it was expected that students in Grades 4 and 8 would find TIMSS assessments too difficult, IEA encouraged countries to test children in higher grades.

¹⁵ Four of these countries are characterised as fragile states (Comoros, Eritrea, Madagascar and South Sudan), which might be one reason for the limited assessment activities.

¹⁶ Zambia will participate in the pilot for PISA for Development in the coming years. This OECD/World Bank-led assessment aims to enhance PISA's survey instruments to make them more relevant for contexts found in developing countries, but still permit the reporting of results on the standard PISA scales. For more information about PISA for Development, see <<http://www.oecd.org/pisa/aboutpisa/pisaforddevelopment.htm>>.

Thus, in TIMSS 2011 Botswana tested children in Grade 6 with the Grade 4 assessment, and children in Grade 9 with the Grade 8 assessment.¹⁷ South Africa took part in TIMSS in 2011 for the first time, testing Grade 9 children with the TIMSS Grade 8 assessment.¹⁸

First introduced in 2001, PIRLS measures trends in reading comprehension at Grade 4. In 2011, PIRLS was expanded to include prePIRLS, which is a less difficult and shorter version of PIRLS. PrePIRLS assesses the basic reading skills at the end of the primary school cycle that are a prerequisite for success in PIRLS (Mullis and Martin, 2013, p. 4). Thus prePIRLS permits learners from lower achieving countries to be measured more precisely than is the case using more difficult and longer assessments, such as PIRLS (Howie, Staden, Tshela, Dowse, and Zimmerman, 2012, p. 22).¹⁹

Participation in regional assessments

Fourteen countries in the region have participated in one or more of the regional assessments of SACMEQ (Southern Africa Consortium for Monitoring Educational Quality), PASEC (Programme for the Analysis of the Educational Systems of CONFEMEN Countries) and Uwezo (see Table 2).

Table 2. Participation in regional assessments in ESAR countries

Country	Assessment
Botswana	SACMEQ II and III
Burundi	PASEC in 2008–2009
Comoros	PASEC in 2008–2009
Kenya	SACMEQ I, II and III Uwezo
Lesotho	SACMEQ II and III
Malawi	SACMEQ I, II and III
Mozambique	SACMEQ II and III
Namibia	SACMEQ I, II and III
South Africa	SACMEQ II and III
Swaziland	SACMEQ II and III
Tanzania	SACMEQ II and III ²⁰ Uwezo
Uganda	SACMEQ II and III Uwezo
Zambia	SACMEQ I, II and III
Zimbabwe	SACMEQ I and III

17 In TIMSS 2007, Botswana participated with Grade 8.

18 TIMSS 2015 introduces a new, less difficult mathematics assessment called TIMSS Numeracy for countries where most children are still developing fundamental mathematics skills. TIMSS Numeracy assesses fundamental mathematical knowledge, procedures, and problem-solving strategies at the end of the primary school cycle that are prerequisites for success on TIMSS (I. V.S. Mullis and Martin, 2013, pp. 7–8).

19 PrePIRLS and TIMSS Numeracy intend to be responsive to the needs of the global education community and efforts to work towards universal learning for all children. Depending on a country's educational development, prePIRLS and TIMSS Numeracy can be given at Grade 4, 5 or 6 (I. V.S. Mullis and Martin, 2013, pp. 7–8).

20 Only Zanzibar of Tanzania participated in SACMEQ I.

SACMEQ carries out large-scale, cross-national research studies in the Southern and Eastern Africa region. It assesses the performance levels of Grade 6 students and teachers in literacy and numeracy (ACER, 2015).

PASEC is an assessment programme for countries that have a link to the French-speaking community. It was established in 1991 by the Conference of Ministers of Education of French-speaking Countries (CONFEMEN), and it assesses Grade 2 and Grade 5 students in reading and mathematics (CONFEMEN, n.d.).

Uwezo measures the literacy and numeracy competencies of school-aged children in Kenya, Tanzania and Uganda. Its goal is to obtain data to inform improvements in educational policy and practice (Twaweza, n.d.).

Implementation of national assessments

In 13 countries in the region, national assessments have been implemented by government or parastatal bodies (see Table 3). In six countries (Eritrea, Ethiopia, Malawi, Rwanda, Somalia and Zimbabwe), an international development partner was directly involved in supporting the implementation.

Table 3. Implementation of national assessments by government/parastatal bodies in ESAR countries

Country	Assessment	Notes
Eritrea	Monitoring Learning Achievement (MLA)	Conducted irregularly by ministry and UNICEF Most recent reported implementation in 2008
Ethiopia	National Learning Assessment (NLA)	Conducted on a 3–4-year cycle by national assessment/examination body and USAID Most recent reported implementation in 2010–2011
Kenya	National Assessment System for Monitoring Learning Outcomes (NASMLA)	Conducted at an unknown frequency by national assessment/examination body Most recent reported implementation in 2010
Lesotho	Lesotho National Assessment of Educational Progress (LNAEP)	Conducted on a 1–2-year cycle by national assessment/examination body Most recent reported implementation in 2010
Malawi*	Assessing Learner Achievement	Conducted by ministry or national assessment/examination body
Malawi	Monitoring Learning Achievement (MLA)	Conducted on an intended 3-year cycle by ministry and UNICEF First implementation in 2012
Mozambique*	National Assessment	Conducted by ministry or national assessment/examination body
Namibia*	National Standardised Achievement Test (NSAT)	Conducted biannually by national assessment/examination body

Rwanda	Learning Achievement in Rwandan Schools (LARS)	Conducted on a 3-year cycle by national assessment/examination body, UNICEF and UNESCO First implementation in 2011, second in 2014
Somalia*	Monitoring Learning Achievement (MLA)	Conducted irregularly by ministry and UNICEF
South Africa	Annual National Assessment	Conducted annually by the ministry Most recent reported implementation in 2014
South Africa*	National Assessment of Learner Achievement (NALA)	Conducted by ministry or national assessment/examination body
Uganda	National Assessment of Progress in Education (NAPE)	Conducted on a 1–3-year cycle by national assessment/examination body Most recent reported implementation in 2010
Zambia	National Assessment of Learning Achievement (NALA)	Conducted on a 2-year cycle by the national assessment/examination body Most recently reported implementation in 2014
Zimbabwe	Zimbabwe Early Learning Assessment (ZELA)	Conducted in 2012–2015 by the national assessment/examination body and UNICEF.

*Note: These assessments are mentioned in the EFA Global Monitoring Report (UNESCO 2008, p. 2; 2015b), but only limited information was available for the main stock-taking table (see Table 24).

Implementation of EGRA/EGMA

The Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA) measure the most basic foundation skills for literacy and numeracy acquisition in the early grades. These assessments were developed by the Research Triangle Institute (RTI), with funding provided by the United States Agency for International Development (USAID) and the World Bank (Gove and Wetterberg, 2011). EGRA/EGMA was designed to serve as a sample-based national or system-level diagnostic measure that would reveal gaps in reading competencies among students and inform education ministries and development partners about system needs for improving the professional development of teachers and pre-service programmes (Gove and Wetterberg, 2011). However, EGRA/EGMA has been used to address a wider range of assessment needs, including impact (programme) evaluations.

Although EGRA and EGMA have been used in many developing countries, including ESA and other regions, they are not grouped as international assessments in this report. This is because EGRA and EGMA have a common approach grounded on core foundation skills. At the same time, they can be adapted for use in individual countries and languages. This approach differs from international assessments where implementing countries are required to use an internationally agreed model (ACER, 2014a). The adaptability of EGRA and EGMA also means that direct comparison of the results is difficult due to differences in language structure and complexity. For this reason, developers of these assessment tools generally advise against comparing subtask results across countries and languages (Gove and Wetterberg, 2011).

Twelve countries in the region have completed at least one implementation of EGRA/EGMA for either system-level diagnostic, monitoring or programme evaluation (see Table 4).²¹

The EGRA/EGMA tools are equally used for both purposes: in nine cases, EGRA and/or EGMA were implemented for system-level diagnostics, and in 10 cases they were used for programme evaluation.²² In two cases, EGRA and/or EGMA were implemented for the purpose of system-level monitoring.

Table 4. Implementations of EGRA/EGMA in ESAR countries²³

Country	Assessment	Year	Purpose
Angola	EGRA	2010	System-level diagnostic
Burundi	EGRA	2011	System level diagnostic
Ethiopia	EGRA	2010	System-level diagnostic
		2010–2012	Programme evaluation (Literacy Boost initiative)
		2011	System-level diagnostic
Kenya	EGRA	2007–2008	Programme evaluation (EMACK initiative)
	EGRA, EGMA	2012–2013	Programme evaluation (PRIMR initiative)
Madagascar	EGRA	2009	System-level diagnostic
Malawi	EGMA	2010	Programme evaluation (baseline for Malawi Teacher Professional Development Support initiative)
	EGRA	2009–2010	Programme evaluation (Literacy Boost initiative)
		2010–2012	System-level monitoring
Mozambique	EGRA	2010–2011	Programme evaluation (Literacy Boost initiative)
		2013	Programme evaluation (USAID/Aprender a Ler (APAL) initiative)
Rwanda	EGRA, EGMA	2011	System-level diagnostic
Somalia	EGRA	2013–2014	Programme evaluation
Tanzania	EGRA, EGMA	2013	System-level monitoring
Uganda	EGRA	2009	System-level diagnostic
		2010, 2012	Programme evaluation (Literacy Boost initiative)
Zambia	EGRA, EGMA	2011	Pilot for system-level diagnostic
	EGRA	2012	Programme evaluation
		2014	System-level diagnostic (as part of the National Assessment Survey)

21 Further details of each EGRA/EGMA implementation can be found in the main stock-taking table in Appendix IV.

22 Assessments with a system-level diagnostic purpose are implemented to get a snapshot of learning levels at the system level (usually national, and in a one-off administration), as compared to assessments with a system-level monitoring purpose which have recurrent administrations to monitor learning levels at the system level.

23 We acknowledge that there are current EGRA/EGMA implementations which are not considered in our stock-taking and that took place between October 2014 and March 2015. Assessments conducted outside this period were not considered. We wish to thank UNICEF CO Rwanda and Tanzania for their input about a School Quality Assessment in three regions using EGRA/EGMA methodology supported by UNICEF which was completed in 2015 focusing on providing baselines in the three UNICEF targeted regions.

1.2 Comparative analysis of assessments

Our comparative analysis of the assessments in the region highlights the strengths and limitations of the different methodological approaches, and discusses when the different approaches might be more or less appropriate. In some instances, best practices in terms of methodological approaches are discussed, even if they do not feature in the majority of the assessments.

The assessments are analysed and compared based on the main eight elements of the stock-taking framework: purpose of the assessment; target population (grade-based, e.g., Grade 4, or age-based, e.g., 10-year-old students); sampling design and methodology; cognitive domains (assessment framework and major domains, i.e., literacy and numeracy); contextual instruments (types, e.g., student questionnaire, and key factors, e.g., gender, grade level, parental education); test administration (approaches for data collection); data analysis (key analytical approaches), and reporting and dissemination products.²⁴

1.2.1 Purpose

There are three types of assessment purposes:

- System-level monitoring: Recurrent administrations to monitor learning levels at the system level (usually national).
- System-level diagnostic: One-off administration for a snapshot of learning levels at the system level (usually national).
- Programme evaluation: Smaller scale administration to evaluate the impact of a programme to improve learning outcomes, with treatment and control groups, and usually involving baseline, mid-line and end-line.

The majority of the assessments in the region (38, or 66 per cent) define their purpose at system-level monitoring. Ten assessments (17 per cent) aim at evaluating programmes. Nine assessments (16 per cent) measure learning outcomes for a system-level diagnostic purpose. One assessment, NASMLA in Kenya, has a dual purpose of system-level diagnostic and monitoring (see Table 5).

Table 5. Purpose of the assessments from stock-taking

Assessment purpose	Number of assessments
System-level monitoring	38
Programme evaluation	10
System-level diagnostic	9
System-level diagnostic and monitoring	1
Total	58

²⁴ Details about the stock-taking framework and other methodological aspects of the stock-taking and comparative analysis are provided in Appendix I.

1.2.2 Target population

The majority of the assessments in the region (52, or 90 per cent) have a grade-based target population. An age-based population is targeted in three Uwezo assessments. There was no information available about the target populations for another three assessments.

For the assessments with grade-based populations, the target populations range from early primary grades to end-of-primary (see Table 6). Of the assessments, 44 target students at lower-primary level (i.e., Grades 1, 2 or 3), 38 target Grades 4, 5 or 6, and 10 target Grades 7, 8 or 9. Nearly half (47 per cent) target multiple grades.

Table 6. Target population in the assessments from stock-taking

Target grade	Number of assessments ²⁵
Grade 1	4
Grade 2	20
Grade 3	20
Grade 4	12
Grade 5	8
Grade 6	18 ²⁶
Grade 7	4
Grade 8	2
Grade 9	4
Age-based	3 ²⁷
Unknown	3

The choice of the grade level for the target population depends on a number of factors. Usually a particular grade is chosen because of recent or planned policy reform, or because it is considered a pivotal point in children's learning trajectories. For example, EGRA and EGMA typically target early primary grades because, as the assessment names indicate ('EG' for Early Grade), they measure the most basic foundation skills for literacy and numeracy acquisition in the early grades (Gove and Wetterberg, 2011).

If the assessment's purpose is programme evaluation rather than system-level diagnostic/monitoring, the grade-based target populations are defined as only children in the treatment or control groups.

For assessments that monitor system-level learning outcomes and report on trends over time, a complete discussion of population-related matters is particularly important, as stakeholders must be informed of any differences that may affect the comparability of results from one year to the next.

²⁵ The total number of assessments in this table is more than 58, as 27 of the 58 assessments target multiple grades.

²⁶ 12 assessments are SACMEQ studies

²⁷ Uwezo in Kenya (6–16 years), Tanzania (6–16 years) and Uganda (7–16 years)

Of the assessments in our stock-taking, the documentation from SACMEQ (Paul M. Wasanga, A. Ogle, and Wambua, 2012), and the IEA studies PIRLS and TIMSS (M.O. Martin and Mullis, 2012) are good examples of a complete discussion of population-related issues, as are the National Learning Assessment in Ethiopia (Ministry of Education of Ethiopia [FDRE], 2008), and NASMLA in Kenya (P.M. Wasanga, Ogle, and Wambua, 2010).

1.2.3 Sampling

In general, the assessments we reviewed employ scientific sampling procedures. The sampling process for assessments with a system-level monitoring or diagnostic purpose is usually multi-step, involving the sampling of schools, children and, sometimes, geographical units.

One exception is South Africa's Annual National Assessment, in which all children in public and state-subsidised independent schools are assessed. This approach may be more accurate, but it is also more time-consuming and costly than testing a sample.

Uwezo, another exception to the general pattern, samples households instead of schools. This may be necessary where a complete and up-to-date list of schools does not exist, but it could make it more difficult to explore the relationships between learning outcomes and school-level factors.²⁸

In four assessments, children were sampled through a central body (e.g., national centre) prior to testing. These assessments included TIMSS, PIRLS (M.O. Martin and Mullis, 2012), LARS in Rwanda (Rwanda Education Board, 2012) and MLA in Malawi (Ministry of Education, Science and Technology of Malawi, 2014). The success of this approach depends upon the availability of complete and up-to-date lists of students before the test is being administered.

1.2.4 Cognitive domains

Assessment framework

An assessment framework is intended to guide test development and help interested stakeholders understand the content and scope of the assessment. In general, a framework should support and ensure consistency of test development and provide a common language for discussing the assessment. An assessment framework should include:

- a definition of the constructs that are being measured;
- a discussion of skills/knowledge that are tested to measure the constructs, as well as provision of a rationale for any omissions of skills/knowledge that one might expect to be tested when the stated constructs are being measured;
- a discussion of any alignment of tests (e.g., to a particular grade level);
- specifications of task content (i.e., number or proportion of tasks per content area);
- specifications of task format (e.g., multiple choice, free-response);
- a discussion of scoring; and
- an outline of how the results are reported.

28 Even though Uwezo's purpose is to monitor children's learning at the system level – like ASER in India, on which its sampling methodology is based – it eschews a more traditional school-based approach to sampling. This is partly because of the difficulty of obtaining a complete and up-to-date list of schools, but also because this approach cannot yield data representative of the entire population of school-aged children in a context where the percentages of school-aged children either not enrolled in school or not attending school regularly are high enough to affect the representativeness of an in-school sample.

If the assessment's purpose is system-level monitoring, and it reports on trends over time, then having an assessment framework is particularly important because it ensures consistency in test development from one assessment cycle to the next.

Comprehensive assessment frameworks were publicly available for three assessments. These include the IEA studies PIRLS (I.V.S. Mullis, Martin, Kennedy, Trong, and Sainsbury, 2009) and TIMSS (I.V.S. Mullis, Martin, Ruddock, O'Sullivan, and Preuschoff, 2009), and the Annual National Assessment in South Africa (Department of Basic Education, Republic of South Africa, 2014).²⁹ PASEC is planning a new methodological framework for the next assessment cycle.

Literacy and numeracy

One-third of the assessments in the region (19, or 33 per cent) assessed both literacy and numeracy.^{30,31} Almost a third (17, or 29 per cent) focused on literacy as the only domain. Twelve assessments (21 per cent, all of which were SACMEQ studies) measured student performance in literacy, numeracy and health knowledge. These three combinations of domains (i.e., literacy and numeracy; literacy only; and literacy, numeracy and health knowledge) constitute the majority of the assessments in our stock-taking, totalling 48 (83 per cent). The rest of the assessments had various combinations of domains (see Table 7).

29 Assessment frameworks for PIRLS, TIMSS and ANA in South Africa are publicly available. Information about the development of an assessment framework for the National Assessment of Learning Achievement in Zambia (UNICEF Zambia Country Office, 2015) was obtained from UNICEF CO in Zambia, but the assessment framework was not publicly available. This may also apply to other assessments in the stock-taking.

30 Domains related to literacy are referred to differently in different assessments. The variations include 'Reading', 'Mother tongue', 'Language', 'English', or the names of particular local languages with slightly different scope of assessment. It is referred as 'Literacy' in this report to avoid confusion. Please see Table 24 for individual cases.

31 'Numeracy' is sometimes referred as 'mathematics' as a domain with slightly different scope of assessment. It is referred to as 'mathematics' in this report to avoid confusion. Please see Table 24 for individual cases.

Table 7. Cognitive domains in assessments from the stock-taking

Domains	Number of assessments	Notes
Literacy and numeracy	19	
Literacy	17	Including PIRLS studies
Literacy, numeracy and health knowledge	12	All are SACMEQ studies
Numeracy and science	2	Both are TIMSS studies
Literacy, numeracy and environmental science (Grade 4)		
Literacy, numeracy, biology, chemistry and physics (Grade 8)	1	NLA in Ethiopia
Literacy, numeracy and life skills	1	Assessing Learner Achievement in Malawi
Literacy, numeracy and life skills (Grade 5)		
Literacy, numeracy and environmental sciences (Grade 9)	1	National Assessment of Learning Achievement in Zambia
Numeracy	1	EGMA in Malawi
Literacy and numeracy (Grade 5)		
Literacy, numeracy and natural science (Grade 7)	1	NSAT in Namibia
Literacy and numeracy (Grade 4)		
Literacy, numeracy and science (Grade 7)	1	MLA in Somalia
Literacy, numeracy and natural science	1	NALA in South Africa
Literacy and one numeracy sub-task	1	EGRA in Uganda
Total	58	

Out of the 58 assessments in the stock-taking, 23 (40 per cent) tested in multiple languages. In this respect, South Africa is the most notable example, as its Annual National Assessment covers English, Afrikaans and nine local languages (Department of Basic Education, Republic of South Africa, 2014).

1.2.5 Contextual instruments

Type of contextual instruments

Almost all of the assessments collect contextual data of some kind through student, teacher and school questionnaires. EGRA and EGMA are often administered using a contextual instrument called the Snapshot of School Management Effectiveness (SSME). The SSME collects information through student and teacher questionnaires, as well as through classroom observation and classroom and school inventories. Uwezo also collects information through a similar technique of observation and inventory. In addition, TIMSS and PIRLS include a curriculum questionnaire completed by the national research centre in each participating country. Furthermore, PIRLS (M.O. Martin and Mullis, 2012), EGRA in Angola (Ministry of Education of Angola, World Bank, and Russia Education Aid for Development Programme [READ], 2011), MLA in Eritrea (UNICEF Eritrea, n.d.) and LARS in Rwanda (Rwanda Education Board, 2012) include a parent

questionnaire. Data on a system level are not directly collected in the assessments, with the exception of the curriculum questionnaire used in PIRLS and TIMSS. However, countries with an Education Management Information System (EMIS) can derive data on a system level, and link it to schools that were observed in the assessment.³²

In assessments undertaken with limited funds, the benefit of obtaining information from a parent questionnaire might be weighed against the risk that response rates will be low due to low literacy skills. An option would be interviews, although these are likely to be more cost-intensive. Experience shows that primary students are a reliable source of information about their parents and households. However, this depends on the type of questions, the kind of data collection instrument (questionnaire or interview), as well as the grade level and associated literacy skills of the children.

Contextual instrumentation must be based on a sound and fully articulated theoretical framework. Good examples for highly elaborated context frameworks are PIRLS and TIMSS. SAQMEC and PASEC use analytical models to describe the context factors collected and the expected relationships with achievement.

Key factors of contextual instruments

Complete contextual instruments (or documentation about its content) were obtained for the following assessments:

- EGRA/EGMA (i.e., SSME instruments) (RTI, 2004)
- PIRLS, TIMSS (IEA, 2013b, 2013d)
- SACMEQ (Hungu, 2011; Hungu et al., 2011)
- Uwezo (Uwezo-Kenya, 2013)
- PASEC (CONFEMEN, 2010a, 2010b)

A review of these instruments and associated documentation suggests that the key factors in contextual data collection on student, classroom and school level are:

- **at the student level:**

- individual factors, such as gender, age, grade level, grade repetition, health and well-being;
- family factors, such as socio-economic measures (possessions at home; books in the home; parental literacy level, education and occupation); ethnic background and cultural practices; language spoken at home; home resources; early learning opportunities (preschool attendance); and family support;
- learning experiences at school (e.g., activities during instruction, teacher feedback provided to students);
- learning experiences out of school (homework and out-of-school lessons; reading independently in and out of school; working outside school/domestic work);
- learning time, attendance/absence;
- access to resources at school; being allowed to take books home;
- community support.

³² In order to link aggregated school data to system level data in EMIS, participating schools need to be able to be identified within the EMIS.

- **at the classroom level**

- teacher background variables, including gender, teacher training, practice and experience class size
- classroom equipment, teaching resources (e.g. availability of pedagogic materials for students and teachers, learning material, text books, furniture)
- language of instruction, language of teacher frequency of homework
- quality of instruction, teaching methods (e.g. teacher reads to learner, explaining things if not understood, providing extra time to complete task; domain-related activities; instructional time), classroom management, frequency and use of assessment for teaching classroom climate, discipline;

- **at the school level:**

- input factors such as school type (public/private); school location (rural/urban); school size; school funding, teacher-student ratio; socio-economic background; and ethnic/language composition;
- teacher body, teacher absenteeism, teacher professional development;
- principal/head teacher background variables; school management, school curriculum, assessment and evaluation; language for instruction, provisions for students who do not speak the language of instruction at home;
- school resources (e.g. library, computer rooms); school facilities (e.g. condition of school building, electricity and water supply, toilets and canteens);
- quality of instruction; school climate.

The collection of contextual information is central to an assessment. Context helps define the relationships between learning outcomes and background factors of research and policy interest. The relationships between contextual factors and achievement – and not the learning outcomes data alone – are essential to decision-making. The time and cost of collecting contextual information can be minimised by ensuring that the instruments are well targeted to specific areas of research and policy interest, and that particular questions yield response data that do not require excessive processing.

1.2.6 Test administration

In just over half of the assessments (30, or 52 per cent), the cognitive assessment is administered as a paper-based test in schools to groups of students, where each student completes the assessment independently (i.e., by reading questions and recording responses on paper). Uwezo, EGRA and EGMA are exceptions. Their cognitive assessments are one-on-one, with the test administrator delivering the items orally, and students providing most of their answers orally. These oral one-on-one assessments generally make use of paper-based instruments, though EGRA and EGMA are starting to employ a method where the student refers to a paper-based test, but the administrator records the data on a tablet-based application called Tangerine™ (see Table 8).³³

³³ See <<http://www.tangerinecentral.org/home>> for information about Tangerine™.

Table 8. Test administration methods of assessments from the stock-taking

Test administration methods	Number of assessments	Assessments
School-based Group administration Paper-based administration	30	PIRLS, prePIRLS TIMSS PASEC SACMEQ National assessments
School-based One-on-one administration Oral administration	20	EGMA EGRA
Household-based One-on-one administration Oral administration	3	Uwezo
School-based One-on-one or small group administration Oral or tablet-based administration	1	Pilot assessment of Grades 1, 2 and 3 in Lesotho
Unknown	4	
Total	58	

A pilot assessment of Grades 1–3 in Lesotho provides an example of how tablets can be used in oral assessments as more than mere data collection tools for the test administrators. In this pilot, students formed a small group with one tablet per student, they interacted directly with the tablets (i.e., the tablets ‘administered the test’ to the students), while the test administrator monitored the groups.

These different methods for cognitive assessments suit different aims and purposes. Group administration is most convenient if members of the target population have the skills to complete an assessment independently, and are easily located in naturally occurring groups (e.g., in schools). One-on-one oral administration is necessary if some students are expected to be unable to complete an assessment independently. If one-on-one oral administration is used, a tablet-based data collection application such as Tangerine™ is an effective way of reducing human error. It can control the way the assessment is administered and restrict data input to only valid-response values. A small group tablet-based oral administration, such as the one used in the Lesotho pilot, offers further efficiency gains because children who would otherwise require one-on-one administration can be tested simultaneously.

1.2.7 Data analysis

The assessments in our stock-taking use a range of different analytical techniques for data analysis. Seven major techniques were identified:

- using item response theory (IRT) to scale cognitive data;
- establishing competency levels or benchmarks;
- conducting frequency analyses or calculating mean scores for cognitive results, disaggregated by contextual variables of interest;

- conducting frequency analyses on contextual data;
- exploring relationships between cognitive performance and contextual factors via analytical techniques;
- computing trends in cognitive performance;
- reporting international comparisons of cognitive data.

Grouping assessments by type reveals patterns of use of particular techniques (see Table 18 in Appendix II). The last row of the table tallies the frequency of use for each technique. Those used most frequently are: establishing competency levels or benchmarks; conducting frequency analyses or calculating mean scores for cognitive results, disaggregated by contextual variables of interest; and exploring relationships between cognitive performance and contextual factors via analytical techniques. These techniques are described in subsequent sections, as is the use of item response theory (IRT).

Use of item response theory

IRT is widely used in the design and analysis of educational assessments. Unlike classical test theory, which is based on the assumption that all items in a test contribute equally to a student's performance, IRT takes into account different characteristics of test items (i.e., varying difficulty levels, discrimination), and hence the probability of getting particular items right or wrong, given the ability of the student taking the test (i.e., the probability that those who do well on the test have a high level of performance, and those who do poorly have low levels of performance) (Kaplan and Saccuzo, 1997). The different approaches account for different levels of test validity.^{34,35,36}

A primary advantage of IRT is that scores obtained from a linked test design – where there are multiple test forms with a certain number of common items shared across the forms – can be placed on a common scale. In IRT, raw scores are converted to scale scores. Placing scores on a common scale permits valid comparisons of results across different test forms, across different grades, and over time.³⁷ Hence, these functions offered by the use of IRT are of particular importance for assessments with system-monitoring purpose.

Another important advantage of using IRT is that it offers a greater depth in reporting. Because item difficulty and ability are on the same scale, it is possible to develop substantive descriptions of the skills and knowledge required to correctly answer items of varying difficulty, and to

34 Classical test theory uses observed test scores of individuals that are composed of a 'true score' (raw score) an individual would get if there were no measurement error. The measurement error, i.e. standard deviation of errors, is assumed to be a random variable with a normal distribution. The larger the standard error of measurement, the less certain is the accuracy of the measurement. Conversely, a small standard error of measurement indicates that an individual score is probably close to the true score (Kaplan and Saccuzo, 1997).

35 Discrimination refers to the capacity of an item to distinguish between different levels of ability (i.e. good quality test questions distinguish between students with the ability to answer the question correctly and those without).

36 For example, one-dimensional models consider item difficulty, two-dimensional models consider item difficulty and item discrimination, and three-dimensional models also account for 'guessing' (i.e. test takers with very low levels of ability getting a correct response).

37 If an assessment does not use a linked design and analyse data using IRT, comparisons of results are only true comparisons if the same assessment items are administered in the same order to all children whose results are being compared. In many cases this is not feasible, as it is often difficult to keep tests secure from one administration to the next. It is also impractical in instances where the construct being assessed is broad, because all items that cover the construct cannot possibly be administered to each child.

make statements about the skills and knowledge possessed by children with different levels of ability. That way, consistent competency levels can be defined and used as a basis for setting benchmarks.³⁸

Moreover, the IRT analysis of the different components provides information on the psychometric quality of the items, e.g., if items are well targeted in terms of difficulty, and can distinguish between students with different levels of ability. With the use of IRT analyses, items of poor psychometric quality can be identified and discarded or adjusted as required.

Of the assessments we examined, IRT is used in the international/regional assessments PIRLS, TIMSS, PASEC and SACMEQ, as well as in five of the national assessments. These assessments have a common purpose: system-level monitoring. However, in a large number of national assessments (seven; the information was not available in four cases) that also aim at system-level monitoring, IRT is not used (see Table 9).³⁹

Table 9. National assessments and use of IRT in data analysis

National assessment	Country	IRT used
MLA	Eritrea	
National Learning Assessment (NLA)	Ethiopia	✓
NASMLA	Kenya	✓
LNAEP	Lesotho	
Assessment of Grades 1, 2 and 3 in Lesotho	Lesotho	✓
Assessing Learner Achievement	Malawi	Unknown
MLA	Malawi	
National Assessment	Mozambique	Unknown
NSAT	Namibia	Unknown
LARS	Rwanda	✓
MLA	Somalia	
Annual National Assessment	South Africa	
NALA	South Africa	Unknown
NAPE	Uganda	
NALA	Zambia	
ZELA	Zimbabwe	✓

In Uwezo and in EGRA/EGMA implementations that aim at system-level monitoring, IRT is not used (see Table 4). This finding shows that despite the advantages of effectively monitoring system performance across contexts and over time, IRT is not extensively applied. One reason for this may be the limited capacity for psychometric analysis at the country level.⁴⁰

38 For some discussion of how results of IRT analysis can be reported and an example of a metric that uses both quantitative and substantive information about children’s performance, see <http://www.acer.edu.au/files/Described_Proficiency_Scales_and_Learning_Metrics.pdf>.

39 A more detailed table showing the analytical techniques used in all assessments from the stock-taking is presented in Appendix II.

40 This was one of the findings of the country case studies conducted in this study (see Appendix III).

Competency levels/benchmarks

In more than half of the assessments (32, or 55 per cent), results are presented with reference to competency levels or benchmarks. In general, competency levels or benchmarks are established by describing the specific skills required to provide correct responses to each test item. Then, test items are placed into groups of items so that the items in each group have similar difficulties and share a common 'theme' relating to the underpinning competencies required to provide the correct response. Naming and defining the 'themes' identifies competency levels or benchmarks (Hungu et al., 2010).

For example, the National Assessments System for Monitoring Learning Outcomes (NASMLA) undertaken in 2010 in Kenya used this analytical technique to assess Grade 3 students in literacy and numeracy. Four competency levels were identified in literacy (The National Assessment Centre, 2010) (see Table 10).

Table 10. Descriptions and attainment of competency levels in literacy: NASMLA in Kenya 2010

Competency level	Description	Percentage of students who attained the level
Level 1	Pre-reading: Matches words and pictures involving concrete concepts and everyday objects.	6.2
Level 2	Emergent reading: Spells correctly simple everyday words and recognises missing letters in such words. Uses familiar words to complete simple everyday sentences.	46.1
Level 3	Basic reading: Uses correct punctuation in simple sentences. Infers meaning from short passages, and interprets meaning by matching words and phrases. Identifies the main themes of a picture.	36.7
Level 4	Reading for meaning: Links and interprets information located in various part of a short passage. Understands and interprets meaning of a picture and writes short sentences to describe the theme.	11.0

Source: Adapted from *Monitoring of Learner Achievement for Class 3 in Literacy and Numeracy in Kenya: Summary of results and recommendations* (The National Assessment Centre, 2010, p. 23).

Students' attainment was analysed as follows: Slightly less than half of the pupils (47.7 per cent) attained the desirable Levels 3 and 4 of competency in literacy. However, most Grade 3 pupils (46.1 per cent) demonstrated emergent reading ability, which is congruent with the Grade 2 level (The National Assessment Centre, 2010, p. 23).

As this example from Kenya demonstrates, competency levels can provide a more concrete understanding of what students are actually able to do than can the insights obtained from merely presenting test scores. Competency levels can also suggest instructional strategies relevant to students who are learning at each level of competence. Such descriptions would be of great

assistance for the preparation of textbooks, the design of teacher in-service training programmes, and the development of general classroom teaching strategies. All of these activities require a sound knowledge of the skills already acquired and the higher order skills that must be mastered in order to move to the next stage of learning (Hungu et al., 2010).

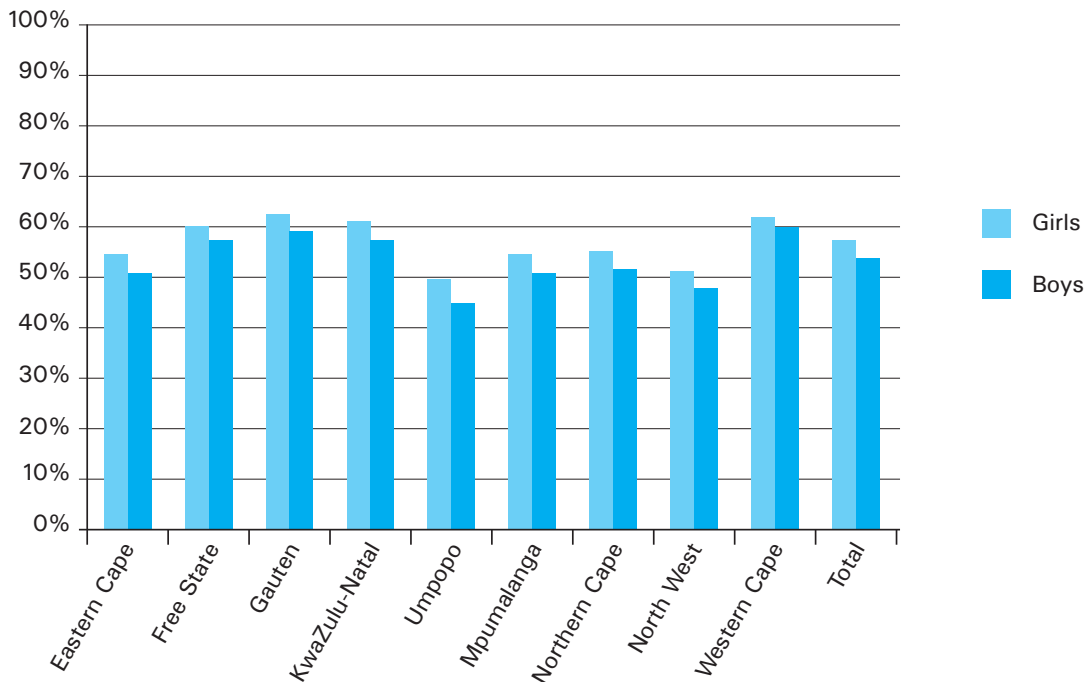
Frequency analysis and mean scores

All of the assessments we studied – other than those without documentation on data analysis – use frequency analysis. These total 50 assessments (86 per cent). Frequency analyses are conducted for performance data and mean scores are calculated. These analyses are usually undertaken on data disaggregated by key contextual variables, such as gender, grade and administrative location.

For example, the Annual National Assessment in South Africa is a system-level monitoring assessment targeting students in Grades 1–6 and 9. It undertakes mean score analysis to investigate the difference in learning achievement between boys and girls (Department of Basic Education, Republic of South Africa, 2014). The mean percentage marks of Grade 3 students in mathematics, calculated by gender and province, show that girls performed better than boys in all provinces (see Figure 3).

Conducting frequency analyses and calculating mean scores are simple procedures, yet they can yield information that is highly relevant to policy development in the implementing countries.

Figure 3. Mean percentage marks in Grade 3 Mathematics by gender and province – Annual National Assessment in Kenya



Source: Department of Basic Education Republic of South Africa, 2014, p. 85

Relationship between cognitive performance and contextual factors

Relationships between cognitive performance and contextual factors are explored in 34 assessments (59 per cent). Analytical techniques used were correlational analysis, regression analysis and multi-level modelling.

EGRA and EGMA in Zambia, implemented in 2011 with support from USAID, provide an example of this approach to data analysis. These reading and mathematics assessments were administered to Grade 2 and Grade 3 students in the Bemba-speaking regions as a pilot study for system-level diagnostic assessment. The final report of the assessments discusses which factors best predict student performance in reading and mathematics. Using multiple regression models, the report shows that five main factors contributed to students' performance in school: socio-economic status; having attended preschool; starting school at the expected age; reading independently in and out of school; and receiving corrective feedback from teachers (Collins et al., 2012).

Conducting analyses such as these helps to ensure that cognitive results are not misinterpreted, as they can be when they are presented without any context. However, the results should not be taken to mean that the relationships are necessarily causal.

1.2.8 Reporting and dissemination

In any assessment programme, availability of quality information and data that address a diverse audience is the key to successful dissemination. Public availability of assessment results and data are important. They allow a wide range of stakeholders to instigate change within the system to improve student performance. The availability of a fully documented database can also, in turn, inform the work of independent researchers.

In the assessments we reviewed, results are publicly available for 41 (71 per cent). Approximately half (22, or 54 per cent) provide additional dissemination products, including results summaries, press releases and policy briefs. Large-scale international and regional assessments, such as TIMSS, PIRLS, SACMEQ and Uwezo, provide ample dissemination products to reach a diverse audience.

Results are not publicly available for 15 (26 per cent) of the assessments we reviewed. For two national assessments where results reports are not publicly available, other means of disseminating them were identified. MLA in Eritrea held workshops at national and sub-national levels (UNICEF Eritrea, n.d.). In Zambia, results of the Grade 5 national assessment in 2008 were disseminated at provincial level. Also, remedial materials were developed for the areas that were found to be challenging for teachers and learners based on the test item analysis (UNICEF Zambia Country Office, 2015).

In addition to reports of the results, considerable effort was made to obtain full datasets from the assessments during our review. However, data from only four regional and international assessments that were implemented in seven ESA countries were available for this study within our time constraints: Uwezo (Kenya, Tanzania, Uganda), PASEC (Burundi, Comoros), TIMSS (Botswana) and prePIRLS (Botswana, South Africa).⁴¹ Data management and data cleaning procedures must be undertaken before assessment data can be analysed, reported and eventually released to the public. The more countries that are involved in an assessment – for example, in international and regional assessments such as PIRLS, TIMSS, PASEC and SACMEQ – the longer it takes for the data to be cleaned at the national and regional/international level and scaled and analysed. Both PIRLS 2011 and TIMSS 2011 studies took approximately two years and three months between main data collection and the release of international datasets (IEA, 2013a, 2013c). Hence, the most recent available datasets for our analysis are for Uwezo 2012, PASEC 2008–2009 and PIRLS/TIMSS 2011 (see Chapter 2).

41 The authors thank UNICEF Headquarters, UNICEF ESARO and COs for their support in requesting access to data from SACMEQ and national assessments. It is acknowledged that in instances where national data were requested from national education ministries or national examination bodies, country-level processes for approving the use of national assessment data for secondary analysis may have required more time than we had available for data analysis. This was the case for national data for South Africa where access was granted outside the time available, and these data were not included.



2. Literacy and numeracy in primary education in the ESA region: Students experiencing LLOs and trends over time

Average test scores for literacy and numeracy in international and regional assessments undertaken in the ESA region were generally low, with a considerable proportion of students not achieving basic skills in reading and mathematics. Results from SACMEQ III (2007) show wide disparities in basic reading and mathematics skills by the end of primary education (Grade 6). In 3 of the 12 participating countries in the ESA region (Kenya, Tanzania and Swaziland) between 80 per cent and 93 per cent of students achieved the minimum reading level in SACMEQ. In six countries (Botswana, Zimbabwe, Namibia, Mozambique, Uganda and South Africa), between 50 per cent and 80 per cent of students achieved the minimum level. In Lesotho, 48 per cent of students in Grade 6 achieved basic reading skills; in Zambia and Malawi, only 27 per cent of students reached this level. In mathematics, the proportion of primary students reaching basic skills was considerably lower, with less than 50 per cent of students in Grade 6 reaching the minimum level in three-quarters of the countries. In the remaining quarter of participating countries (again Kenya, Tanzania and Swaziland) between 56 per cent and 62 per cent of students learned basic mathematics skills (UNESCO, 2014, p. 35).⁴²

Characteristics of low-performing students and trends in student performance over time in literacy and numeracy in primary education are the focus of this chapter. Specific analyses draw upon data from four different assessments in the region (as discussed in Chapter 1): Uwezo (Kenya, Tanzania and Uganda); PASEC (Burundi and Comoros); prePIRLS (South Africa and Botswana); and TIMSS (Botswana).⁴³ The four assessments cover 7 of the 21 ESA countries. Trends in literacy and numeracy performance were analysed for three countries – Kenya, Tanzania and Uganda – where the same assessment, Uwezo, with the same key features (assessment framework, design, target population and conditions of administration) was implemented more than once.⁴⁴ Data from the regional assessment SACMEQ, in which 12 ESA countries participated at least twice, and data from national assessments, were not available for this study.⁴⁵ The limited available data makes it difficult to draw conclusions about the characteristics of low-

42 Source: IIEP/Pôle de Dakar Indicator Database (UNESCO, 2014, p. 35); (Hungu et al., 2010)

43 For details about the datasets available, see Appendix I.

44 Any comparison of results between different assessments—or between different cycles of the same assessment that don't have the same key features across the different cycles—requires sophisticated linking procedures and analyses that are beyond the scope of this report.

45 The authors thank UNICEF Headquarters, UNICEF ESARO, and UNICEF COs for their support in requesting access to data from SACMEQ and national assessments during the time of this study. Findings from SACMEQ and national assessments are referred to in the discussion where reports including analysis of contextual data were obtained.

performing students and trends in performance over time for the region. Given these constraints, we drew upon findings from a broad variety of reports from the ESA region, including SACMEQ and national assessments, to supplement our analyses. The findings that were reported provide valuable insights into the main characteristics of students experiencing LLOs in literacy and numeracy, the nature of associations between context factors and student performance, and changes in student performance over time.

2.1 Characteristics of students experiencing LLOs in literacy and numeracy in primary education in the ESA region

Equipping all students with basic literacy and numeracy skills, and minimizing the number of low-performing students in these domains, are fundamental goals of education systems. In order to quantify the number of students at different levels of performance and to monitor progress over time, competency levels must be defined and benchmarks set. Understanding the factors associated with low-performing students is critical for the development of targeted education policies.

As outlined in the previous chapter, 32 out of the 58 assessments that we reviewed define competency levels or benchmarks. However, there are no common metrics in literacy and numeracy across the different assessments, and different benchmarks are used to define 'limited learning outcomes'. Hence, for our study limited learning outcomes are based on the benchmarks used for literacy and numeracy in the different assessments analysed in this report: PASEC, Uwezo, TIMSS and prePIRLS. The criteria used to set these benchmarks are conceptually quite different.⁴⁶

For PASEC (Burundi and Comoros), student achievement scores in reading (French, Kirundi) and mathematics are categorised in three levels.⁴⁷ Students at Level 3, the highest level, have acquired a basic level of knowledge (CONFEMEN, 2010b, p. 93). At the other end, at Level 1, students are considered to be close to failing. This group is categorised as experiencing LLOs.⁴⁸

Following Uwezo's approach, in which tests are aligned with the national Grade 2 curriculum in the three participating countries (Kenya, Tanzania and Uganda), all children attending Grade 3 are expected to have achieved the highest level in each domain ('story' for literacy and 'multiplication/division' for numeracy). Hence, children enrolled in Grade 3 and above who have not achieved the highest level of performance in English, Swahili and numeracy are considered to be experiencing LLOs.

⁴⁶ Details about the definition and identification of students with limited learning outcomes in the different assessments are presented in Appendix I.

⁴⁷ Grade 2 students from Burundi were also assessed in Kirundi.

⁴⁸ Swahili was only assessed in Kenya and Tanzania.

For TIMSS (Botswana) and PIRLS (or prePIRLS) (Botswana and South Africa) four international benchmarks are defined, ranging from 'advanced' to 'low' (I. Mullis, Martin, Foy, and Arora, 2012; I. Mullis, Martin, Foy, and Drucker, 2012).^{49,50} Students reaching the 'Low International Benchmark' for Grade 4 show some basic mathematical knowledge and reading skills. For the purposes of our analysis, students who have not achieved the Low International Benchmark are considered to have experienced LLOs.

We listed in a table the percentage of students experiencing LLOs, organised by assessment, country, grade and domain, as well as the criteria used for defining the LLOs for the different assessments (see Table 11).

Due to the conceptually different criteria or benchmarks used, the percentage of students experiencing LLOs varies considerably across the assessments. In Burundi and Comoros (PASEC), approximately one in five students experienced LLOs, with a similar percentage for literacy and numeracy domains across Grade 2 and Grade 5. In Kenya, Tanzania and Uganda (Uwezo), the percentage of primary school age students experiencing LLOs differed between the domains. For mathematics, approximately one-third of the students experienced LLOs. For literacy, the percentages varied across the Uwezo countries. In Kenya, around one in five students performed low in English or Swahili. In Tanzania, every second student of primary school age experienced LLOs in English, and approximately every third student in Swahili. In Uganda, 39 per cent of primary school age students experienced LLOs in English. In Botswana (TIMSS), 40 per cent of Grade 6 students experienced LLOs in mathematics. Around one in four students in Botswana and South Africa showed LLOs in reading (prePIRLS).⁵¹

49 Botswana participated in TIMSS 2011 with Grade 6 students, using Grade 4 TIMSS assessment. If it was expected that a country's Grade 4 students would find TIMSS assessment too difficult, IEA encouraged the country to test children in a higher grade.

50 PrePIRLS was chosen over the traditional PIRLS dataset as it is better targeted towards the achievement of students from participating countries for the region. Botswana and South Africa participated in prePIRLS with Grade 4 students in 2011.

51 It is worth noting that Grade 6 students in Botswana were assessed using the TIMSS test material targeted to Grade 4; the proportion of low performing students in mathematics would presumably be higher if Grade 4 students were tested with the Grade 4 assessment material. For reading, Grade 4 students in Botswana and South Africa were assessed with the 'easier' or better targeted prePIRLS test materials; presumably the proportion of low-performing students would be higher if measured with the standard PIRLS tests.

Table 11. Proportions of students defined as experiencing limited learning outcomes by assessment, country, grade and domain

Assessment (Year in brackets)	Criteria for limited learning outcomes	Country	Grade	Mathematics literacy	Reading literacy
PASEC (2008/2009)	Students with a test score of less than 25 out of 100	Burundi	Grade 2	18%	19% (French), 18% (Kirundi)
		Burundi	Grade 5	23%	22%
		Comoros	Grade 2	21%	21%
		Comoros	Grade 5	24%	20%
UWEZO (2012)	Students enrolled in Grade 3 and above who could not achieve the highest level of performance	Kenya	Primary school age	39%	23% (English), 21% (Swahili)
		Tanzania	Primary school age	31%	50% (English), 32% (Swahili)
		Uganda	Primary school age	32%	39%
TIMSS (2011)	Students scoring below 'low achievement threshold' proficiency standard (score of less than 400)	Botswana	Grade 6	40%	
Pre-PIRLS (2011)	Students scoring below 'low achievement threshold' proficiency standard (score of less than 400)	Botswana	Grade 4		23%
		South Africa	Grade 4		29%

2.1.1 Individual and family characteristics of students with LLOs in literacy and numeracy

Individual and family characteristics of students found to be important in association with LLOs are: gender; age; language spoken at home; a range of socio-economic factors; learning activities prior to attending school; engagement in reading lessons; and whether the student attends lessons out of school. Overview tables with detailed results for each of the factors observed in the analysis per country, grade and domain are presented in Appendix IV.

Gender

In general, the assessments showed that males were more likely to experience LLOs in literacy than females. This pattern was found in Burundi (PASEC, both French and Kirundi), Kenya (Uwezo, both English and Swahili), South Africa and Botswana (prePIRLS, English). The most extreme example was in Botswana, where males were almost three times more likely than females to be experiencing LLOs. The exception was Comoros, where females from Grade 5 were slightly more likely to be experiencing LLOs in French than males (1.1 times more likely). It is not clear why the direction of the gender difference was reversed in Comoros, but we assume other factors contributed to it. For example, females in Comoros with illiterate parents were noticeably more likely to experience LLOs than males.

The findings for gender and its relationship with literacy mirror previous findings from the region. Girls were reported to outperform boys in reading literacy in Botswana (Monyaku, 2012); South Africa (Moloi and Chetty, 2010); Zimbabwe (ACER and ZIMSEC, 2015); and Eritrea (UNICEF Eritrea, n.d.). In contrast, an EGRA study in Ethiopia showed that Ethiopian boys had higher levels of early reading scores than girls. However, there was an interactional regional effect. Rural boys outperformed rural girls on almost all tasks, but in urban schools, the opposite effect was the case, with girls outperforming boys (RTI, 2010, p. 37).

For mathematics, girls were found to have greater representation in the LLO group in Burundi and Comoros, but less in Kenya and Botswana. Again, the differences across countries are likely attributable to other factors. For example, in Botswana, a greater percentage of students considered relatively old for their grade (aged above 14 years) were male compared to other age groups.

The different directions of the gender differences found in the assessments are consistent with the findings in prior studies from the region. Boys were reported to have outperformed girls in mathematics in Kenya, Uganda and in provinces from Somalia (Report on Monitoring Learning Achievements (MLA) in Grade 4 in Puntland and Somaliland, 2012; Uganda National Examinations Board, 2010; Paul M. Wasanga et al., 2012). However, the opposite pattern was found in Zimbabwe and South Africa (ACER and ZIMSEC, 2015; Moloi and Chetty, 2010).

Age

Age – relative to school entry and grade – is an important factor associated with student performance in the ESA region. However, the relationship between age and performance is complex, and may be influenced by other factors, such whether school entry occurs at the official starting age, the student’s cognitive development at the time of school entry, prior learning opportunities that affect progression through the grades, as well as instructional practices in

response to student diversity.⁵² A literature review conducted for a study by Hungi et al. (2014) found different effects of age – relative to the grade where the students are at – for developed countries, where ‘older’ students in class generally outperformed their ‘younger’ colleagues, and developing countries, where most studies, especially from Africa, show that ‘younger’ students perform better than ‘older’ students (Hungi et al., 2014, p. 249).⁵³

The assessments we reviewed showed that across the region younger students consistently performed better.⁵⁴ For example, Grade 6 students in Botswana who were 12 years or under (approximately one-fifth of the population) were almost three times less likely to be experiencing LLOs in mathematics than students aged over 12 years of age. This is well supported by the literature in the region. A study by Kunje (cited in Hungi et al., 2014) showed that younger students in Grade 7 in Malawi outperformed their older colleagues in English literacy, Chichewa (a local language) and mathematics. SACMEQ III in 2007 showed that in 12 out of 15 countries, younger students performed better than older students; in nine of those countries, younger students performed better in reading as well as in mathematics (Hungi et al., 2014). Additionally, data from 740,000 students from the 2010 Kenya Certificate of Primary Education examination showed that younger students performed better than older students, according to Keith et al. (cited in Hungi et al., 2014).

Students relatively older than the average class age were over-represented in the LLOs groups for Botswana (prePIRLS, TIMSS), South Africa (prePIRLS) and Grade 5 students in Comoros and Burundi (PASEC). In contrast, Grade 2 students in Burundi who were relatively older, and older children from Kenya, were less likely to be experiencing limited learning. These mixed findings may be attributable to the different year-levels examined. Students relatively older in the latter years of primary school would have been more likely to have repeated a grade than students in the earlier years. At earlier years, age differences are more likely due to different ages of school commencement. For example, of the students from Grade 5 in Comoros who were considered relatively older, two-thirds had repeated a grade at some stage. These students were considerably more likely to be experiencing LLOs than those who did not repeat (and therefore started school at a later age).

The 2012 Monitoring Learning Achievement project (MLA) for Malawi found that those learners who had repeated at least one grade scored significantly lower than those who had never repeated a grade. The achievement levels of those who repeated were generally low and were more evident in higher than lower grades (Ministry of Education, Science and Technology of Malawi, 2014). These studies suggest that while grade repetition is not necessarily the cause of poor performance it does indicate that repeating a grade may not help low-performing students, regardless of age (Njora Hungi et al., 2014).

52 For example, as reported in Hungi, Ngware, and Abuya (2014), reasons given for early school entry by some parents in Kenya is ‘... a hope, that, if the children do not do well, they can always repeat because they have a year or two to spare compared to their classmates’ (Hungi et al., 2014, p. 256).

53 Hungi et al. (2014) investigated the optimal age with the greatest positive impact on literacy achievement for Grade 6 students from low-income families across six major slums in Kenya. For this study, a sample of 7041 Grade 6 students from 226 schools across six major urban slums in Kenya was drawn (Hungi et al., 2014, p. 247).

54 The criteria and relative proportions in each age group varied across countries and datasets. In PASEC for Burundi and Comoros this was defined as 5 years old or below for Grade 2 students and 8 years or below for Grade 5 students. For prePIRLS for South Africa and Botswana (also for TIMSS) this was defined as 12 years or less. For Uwezo for Kenya, Tanzania and Uganda this was defined as children who were aged between 6 and 9 years. From the data we are unable to determine whether younger students began school at an earlier age or whether a high proportion of other students were late entrants to schooling or had repeated a grade.

Language spoken at home

The language spoken at home, or rather the degree of alignment between the language spoken at home and the language of instruction, has a strong impact on learning outcomes. In multilingual environments, such as the ESA region, options for multiple language instruction are evident, which can pose complex challenges to education policy and management (Heugh, Bogale, Benson, and Yohannes, 2006). Research on first language instruction shows that children benefit from mother-tongue instruction for their cognitive development in general, and early literacy acquisition in particular (Ball, 2010; Bialystok, 2001; Cummins, 2000; Heugh et al., 2006; RTI, 2010; UNESCO, 2010).⁵⁵ However, the alignment between language spoken at home and language of instruction is just one factor affecting student achievement. Other interacting factors are socio-economic status; the overall quality of language/reading instruction and instruction in general; provision and use of language instruction materials/books; linguistic complexity; the teacher's proficiency in the language of instruction; teacher training; and the school environment. Given these complex interactions and particular country contexts, further research is needed on language of instruction policies and practices (RTI, 2010).

The language spoken at home was captured for Burundi, Comoros (both PASEC), South Africa (prePIRLS) and Botswana (TIMSS and prePIRLS). In most instances across the assessments, speaking the test language at home was an advantage for children. In Comoros, where students were assessed in French, 97 per cent of all Grade 2 students spoke Shikomori at home (96 per cent of Grade 5 students), while only 3 per cent spoke French at home (4 per cent of Grade 5 students). Those students who spoke French at home were less likely to experience LLOs in both French and also in mathematics. In Burundi, 95 per cent of students at Grade 2 level speak Kirundi at home (which is the teaching language until Grade 4), but were no more likely to experience LLOs for this language; however, they were over-represented in the LLOs group for French language assessment. For South Africa in prePIRLS, the vast majority (91 per cent) of students spoke the language of the assessment at home at least sometimes.⁵⁶ For Botswana, in comparison, approximately three in four students spoke the language of assessment (English) at home (74 per cent in prePIRLS, 78 per cent in TIMSS). Those who spoke the test language at home were less likely to be experiencing LLOs in South Africa, with mixed findings found for Botswana (a difference was found in TIMSS but not for prePIRLS). It is not apparent why the effect of test language was not consistent across studies. However, the data does suggest that any relationship between speaking the language of test at home and performance is likely moderated by the influence of socio-economic factors. Students from both South Africa and Botswana who always (or almost always) spoke the language of test at home were more likely to have greater home resources.

⁵⁵ A child's first language is also often referred to as mother tongue and as the language spoken at home in assessments.

⁵⁶ In South Africa, the prePIRLS assessment was administered in the 11 official languages, and more than 90 per cent of the population were administered the test in a language they speak at least sometimes at home. Information from the South African PIRLS 2011 national report indicates that less than 10 per cent of the population mainly speak English at home, despite English being the language of instruction for almost 80 per cent of the population at Grade 4 (Howie et al., 2012).

Other studies within the ESA region show similar results. Evidence from PASEC and SACMEQ show a strong link between home language and the language of instruction in influencing test scores (Fehrler and Michaelowa, 2009, cited in UNESCO, 2010, p. 154; Garrouste, 2011). While this is more commonly seen as a factor in literacy assessments, evidence from Namibia using SACMEQ results suggests that low language skills also are a large contributor to low performance levels in mathematics (Garrouste, 2011, p. 231). Students from Zimbabwe's ZELA project who spoke English at home had greater performance on the English and mathematics tests than students who spoke other languages at home (the mathematics test was also administered in English).⁵⁷ Very few students (3 per cent) spoke English at home, as opposed to the main language spoken at home, which was Shona (70 per cent) (ACER and ZIMSEC, 2015, p. 29).

Socio-economic factors

The socio-economic status of students is a strong predictor of achievement. The relationship between socio-economic factors and achievement exists consistently across the different assessments. Our analyses support this finding: students from lower socio-economic backgrounds were more likely to experience LLOs across all countries examined in both literacy and numeracy. This relationship was found across all assessments examined, even when the assessments (and the countries within the assessments) used different measures of socio-economic status.⁵⁸

Parental education and literacy was also consistently found to be associated with LLOs. The mothers of students from Kenya, Tanzania, Uganda, South Africa and Botswana experiencing LLOs were more likely to have no formal education or lower levels of education (the same pattern was found for fathers for the latter two countries).⁵⁹ Similarly in Burundi and Comoros, the most common pattern was that students identified with LLOs were more likely to have illiterate mothers and fathers.

Other socio-economic measures we examined showed that the socio-economic background of students had an effect on their school performance. A number of indicators for home possessions were identified, such as housing materials. In Kenya and Tanzania, students were more likely to experience LLOs (literacy and mathematics; Uwezo) if they indicated that the walls of their houses were made of the least expensive material of the possible options (mud) and less likely to experience LLOs if the walls were made of the most expensive materials of the options (stone/brick). This is not to suggest that there is a direct link between housing material and student learning but rather using housing material as a proxy for socio-economic status supports the finding that student performance is affected by home background.

57 The ZELA English test was administered in the context of the national policy that children in Zimbabwe learn English as a subject from Grade 1 (ACER and ZIMSEC, 2015, p. 36).

58 The assessments analysed as part of this study did not have data related to nutrition, child health or other related child development factors, which may constitute an area of further investigation with regard to student performance.

59 Uwezo only collected data on maternal education for Kenya, Tanzania and Uganda, whereas TIMSS and prePIRLS collected data for both parents.

The assessments we analysed all included measures of household possessions. Home possession indices for Comoros and Burundi (PASEC) and for Botswana (TIMSS and prePIRLS) and South Africa (prePIRLS) in general showed that students experiencing LLOs, regardless of the domain, were more likely to have fewer household resources.^{60,61} In general, across all assessments, students experiencing LLOs were less likely, compared to their peers, to have a range of household possessions (e.g., fridge, TV, phone, books) basic facilities (access to electricity or clean water) and individual possessions such as children’s books.

Students experiencing LLOs in Comoros were also found to have more outside work activities than other students, including farm, household and retail work. These students were also more likely to indicate that their work hindered their ability to study at home, attend school and concentrate at school. The same association between non-school-related work and students was not found for students from Burundi. A similar association was found in ZELA, where the amount of time that students spent working was negatively associated with achievement data (ACER and ZIMSEC, 2015).

Learning activities prior to attending school

Analysing data from Botswana and South Africa (prePIRLS and TIMSS) we found that students who attended preschool (46 per cent of Grade 4 students in Botswana and 83 per cent in South Africa) were less likely to experience LLOs than students who did not.⁶² This is a well-supported finding. Other regional studies from Zimbabwe, Zambia and Malawi found a relationship between preschool attendance and achievement levels (ACER and ZIMSEC, 2015; Collins et al., 2012; Ministry of Education, Science and Technology of Malawi, 2014). However, it is important to note that the relationship is not necessarily causal. For example, students who attended preschool in Botswana and South Africa were also more likely to have greater home resources for learning and higher levels of parental education.

Parents of students in Botswana and South Africa were also asked about their children’s exposure to activities prior to attending school related to reading (prePIRLS) and mathematics (TIMSS), as well as their level of competency in these domains once they started school. Early activities included reading books, telling stories, singing songs, playing word games, writing letters or words, reading aloud signals and labels. Examples of early numeracy activities are counting different objects, playing games involving shapes, playing with building blocks or construction toys, or playing board games or card games. Students experiencing LLOs were less likely to have had exposure to such activities before attending school and were rated as having lower levels of competency at commencement. Again, the socio-economic status of the family was found to be associated with these two measures. This suggests that families with greater home resources are more likely to engage in learning activities with their children, which, in turn, is likely to increase the capabilities of students when commencing school.

60 This index was based on the following home possessions and basic facilities: electricity, a television, a telephone, a fridge, gas heating, a video recorder, a computer and a car.

61 This index includes number of books in the home, number of children’s books in the home, number of home study supports, highest parental education level, highest parental occupation level.

62 45 per cent of Grade 6 students in Botswana assessed in TIMSS attended preschool.

Engagement in reading lessons and out-of-school lessons

Students from South Africa and Botswana (prePIRLS) were presented with a series of statements that probed their engagement with reading lessons.⁶³ Their responses were categorised as 'Engaged', 'Somewhat engaged' or 'Not engaged'. Students who were 'Engaged' were far less likely to be experiencing LLOs. Those who were 'Somewhat engaged' or 'Not engaged' were much more likely to be experiencing LLOs. In South Africa, students who fell into the latter category were almost three times more likely to be experiencing LLOs. It is important to note that the relationship between engagement and performance is likely to be reciprocal: more engaged students are more likely to perform better and higher performing students are more likely to be more engaged.

The indicator of home resources for learning was lower for those students who experienced LLOs and were categorised as 'Not Engaged', suggesting that socio-economic status may influence engagement. Howie and colleagues emphasize the need for schools to provide students with a 'variety of stimulating, developmentally appropriate reading materials aligned to teaching practices that encourage active learning on the part of learners', in order to help them engage with their reading (Howie et al., 2012, p. 109).

Kenyan students who were found to be experiencing LLOs (English, Swahili or mathematics) were less likely to have received extra lessons or tuition. Much of this may be explained by other demographic factors. For instance, students who received extra lessons or tuition were more likely to have attended private schools, have a higher number of household possessions and have greater access to basic facilities.⁶⁴ Parents in Kenya, as well as Mauritius, have been reported to be among the highest users of extra tuition outside school (Paviot et al., 2008). However, it may be that the relative costs of such services mean they aren't often used by families with fewer resources, regardless of academic need (Buchman, 2000).

2.1.2 School-level characteristics of students with LLOs in literacy and numeracy in ESAR

School-level characteristics that were observed to be factors related to LLOs are school resources, school type and school location. Overview tables with the detailed results for each of the factors observed in the analysis per country, grade and domain are presented in Appendix IV.

School resources

School resources captured in PASEC, prePIRLS and TIMSS included in this analysis cover electricity, drinking water facilities, toilets, a school library, and a computer room or computer for instruction. The proportion of students attending schools where such facilities are available varies broadly. For example, a maximum of 27 per cent of Grade 5 students in Comoros attended a school with electricity, as opposed to 7 per cent in Burundi. Students in Comoros also had more access to drinking water in school (71 per cent versus 40 per cent of Grade 5 students in

63 Items that form the scale include 'I like what I read about in school'; 'My teacher gives me interesting things to read'; 'I know what my teacher expects me to do'; 'I think of things not related to the lesson'; 'My teacher is easy to understand'; 'I am interested in what my teacher says'; 'My teacher gives me interesting things to do'.

64 The available data for this study do not provide information about the quality of the extra lessons and their direct impact on learning outcomes.

Burundi). Toilets were available for at least three-quarters of students in Comoros, and for nearly all students (92 per cent) in Burundi. A school library was available for a maximum of 12 per cent of students in Grade 5 in Comoros and 3 per cent of Grade 2 students in Burundi. Approximately half of the students participating in prePIRLS and TIMSS in South Africa and Botswana were in schools with a library. The biggest difference was found for computer resources: while 1 per cent of students in Comoros and Burundi were in schools with a computer room, between 48 per cent of Grade 4 students in South Africa and 70 per cent of Grade 6 students in Botswana attended a school with computers for instruction.

Students who attended schools with access to electricity and drinking water facilities in Comoros and Burundi (PASEC) were less likely to experience LLOs than students attending schools without such facilities. The ZELA study in Zimbabwe also found adequate water and electricity resources to be strongly associated with student achievement even once other factors such as home resources of the student are taken into account (ACER and ZIMSEC, 2015).⁶⁵

In Botswana (TIMSS, prePIRLS) and South Africa (prePIRLS), smaller percentages of students in schools with library facilities than those without experienced LLOs. This pattern was not observed in Comoros and Burundi (PASEC), where far fewer schools had access to this resource. Similar results were observed with access to school computers. Schools in Botswana and South Africa with computers for instruction tended to have fewer students experiencing LLOs, whereas the proportion of schools in Comoros and Burundi with computer rooms was negligible.

Majgaard and Mingat (2012) provide a comprehensive overview of school inputs that contribute to learning achievement in primary schools in low-income sub-Saharan African countries. In their report, the authors combined test scores from three international learning assessment programmes – SACMEQ, PASEC and MLA surveys – to create a comparable Africa Student Learning Index (ASLI) for the region. Their major findings at the school level on how learning outcomes can be improved include observable characteristics such as the quality of school buildings and the availability of libraries, although the authors note that it is likely that resources by themselves do not necessarily improve learning. Pedagogy plays a large role. Studies that we analysed show that school resources such as clean water, adequate sanitation and access to suitable learning and reading materials, such as the provision of libraries, are associated with positive student learning outcomes.

School type and school location

Data from Kenya, Tanzania and Uganda (Uwezo) indicate an association between school type (public versus private) and student outcomes. Students attending public schools were more likely to be experiencing LLOs than those attending private schools.

In Botswana (TIMSS, prePIRLS) and South Africa (prePIRLS), students attending rural schools were more likely to be experiencing LLOs than their peers attending urban schools.

School type has consistently been shown as a good predictor of achievement (ACER and ZIMSEC, 2015; Ministry of Education, Science, and Technology of Malawi, 2011; Wasanga et al., 2012). Students attending non-government schools traditionally outperform those attending

⁶⁵ The study reported the presence of electricity and water facilities at the school to be significantly associated with mathematics achievement for both rural and urban areas, but only rural (and not urban) areas for English achievement.

government schools (Uganda National Examinations Board, 2010). Non-government schools also did better than government schools in the three learning areas (UNICEF Eritrea, n.d.). The same pattern has been reported in numerous studies regarding school location. Students attending schools in urban environments tend to have greater achievement levels than those attending schools in rural environments, a finding almost universally consistent across studies (ACER and ZIMSEC, 2015; Makuwa, 2005; Ministry of Education et al., 2011; Moloi and Chetty, 2010; Wasanga et al., 2012).

This is not to say that schools perform better because they are located in urban areas. The amount of resourcing that a school has available likely explains the relationship. As an example, data from Botswana (TIMSS) show that urban schools were more likely to be private and better resourced than schools in rural areas. Indeed, it is likely that school resourcing accounts for much of the variation found between different school types, and between different school locations.

2.2 Trends in learning outcomes of children in primary education in literacy and numeracy in the ESA region

Tracking progress in student performance over time is an important element of system-level monitoring in education. As discussed in Chapter 1, approximately two-thirds of the assessments we reviewed have a system-level monitoring purpose, and conduct recurring assessments to monitor changes in students' literacy and numeracy performance levels. Apart from SACMEQ, PASEC, Uwezo and PIRLS/TIMSS, these assessments are conducted at a national level. As a regional assessment in Southern and Eastern Africa, SACMEQ is best placed to provide comparable data for a large number of countries (12 ESA countries participated at least twice). Since SACMEQ data were unavailable for this study, trends in performance were analysed for the three countries participating in Uwezo: Kenya, Tanzania and Uganda.⁶⁶

2.2.1 Uwezo literacy trends

Performance data for Uwezo relate to the percentage of primary-school-aged students that are able to successfully complete each of the tasks within a learning domain.⁶⁷ Hence performance is presented at task level for each participating country for each year to explore changes over time.

Students undertaking the literacy component of Uwezo are rated on their highest level of task completion on this ascending scale of difficulty: 'nothing', 'syllables', 'words', 'paragraphs' and 'story'. We show in the bar graph below an example of performance levels for Kenyan students completing the English literacy component over the three years for which data are available (see Figure 4). The data are presented in a hierarchical format that displays the highest task that each student could complete. For 2009–2010, it shows that 49 per cent of students were able to complete the story task. As this is the most difficult, this group of students would therefore also have successfully completed the paragraph, word and letters tasks. Only 6 per cent of students were not able to complete the most basic task (letters).

⁶⁶ This study in particular looks at trends in performance since 2007. At the time of data analysis, PASEC was last implemented in Burundi and Comoros in 2008/2009 (only once). The first administration of prePIRLS was in 2011. Botswana had participated in TIMSS before, but with varying target populations (TIMSS 2007: Grade 8; TIMSS 2011: Grade 6 and Grade 9).

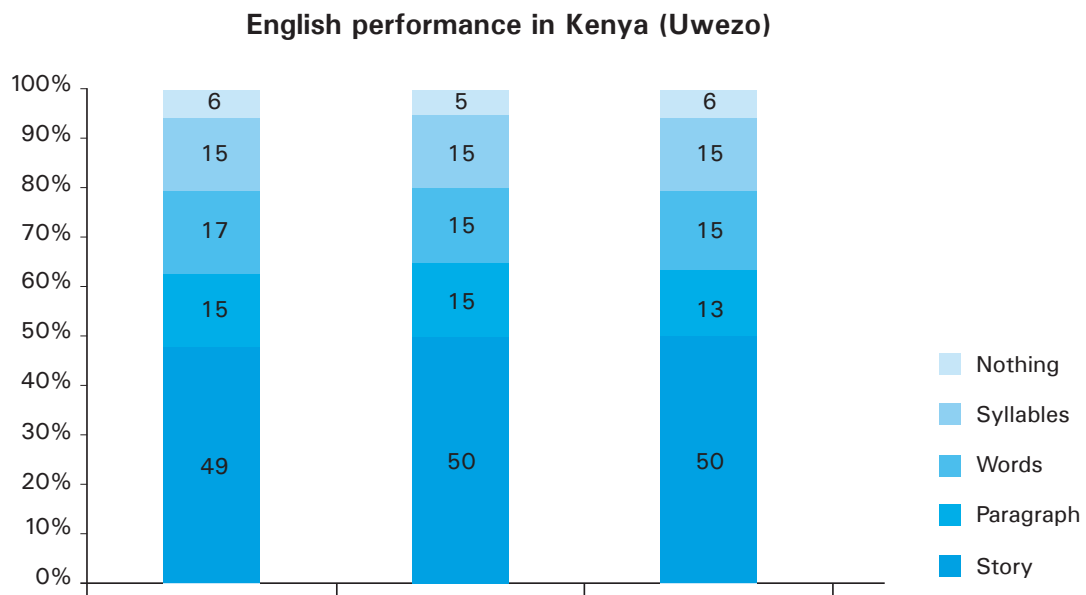
⁶⁷ See Appendix I for further discussion of Uwezo data-related issues for trend analyses.

Figure 4 shows a similar proportion of students being able to complete each task across the three years. The differences in proportions across time for each task – or changes in student performance – are all relatively minor. Similar figures for trends in English are provided in Appendix IV for Tanzania and Uganda, and for Swahili in Kenya and Tanzania.

In Tanzania, there was little improvement over time for English in terms of the proportion of students able to complete the most complicated task (story). However, there was a significant reduction in those not able to complete any task, dropping from 34 per cent in 2009–10 to 29 per cent in 2012. This difference was still significant – although reduced in magnitude – after a model that incorporated age and gender was introduced. In Uganda, the opposite pattern was found. There was little change in the proportions of students who could not complete any task, but a significant increase in the proportion of students able to complete the most complicated task, rising from 17 per cent in 2009–10 to 27 per cent in 2012. Because the populations of students that completed the assessments were different, a separate analysis was conducted to determine whether age and gender might account for the difference. Even after these two demographic variables were controlled for, the increase remained significant.

For Swahili, there were few observable trends found across Kenya, but in Tanzania there was a noticeable drop in the proportion of children able to complete the most difficult task, from 45 per cent in 2009–10 to 36 per cent in 2012. The drop remained significant even after the age and gender of the populations were taken into account.

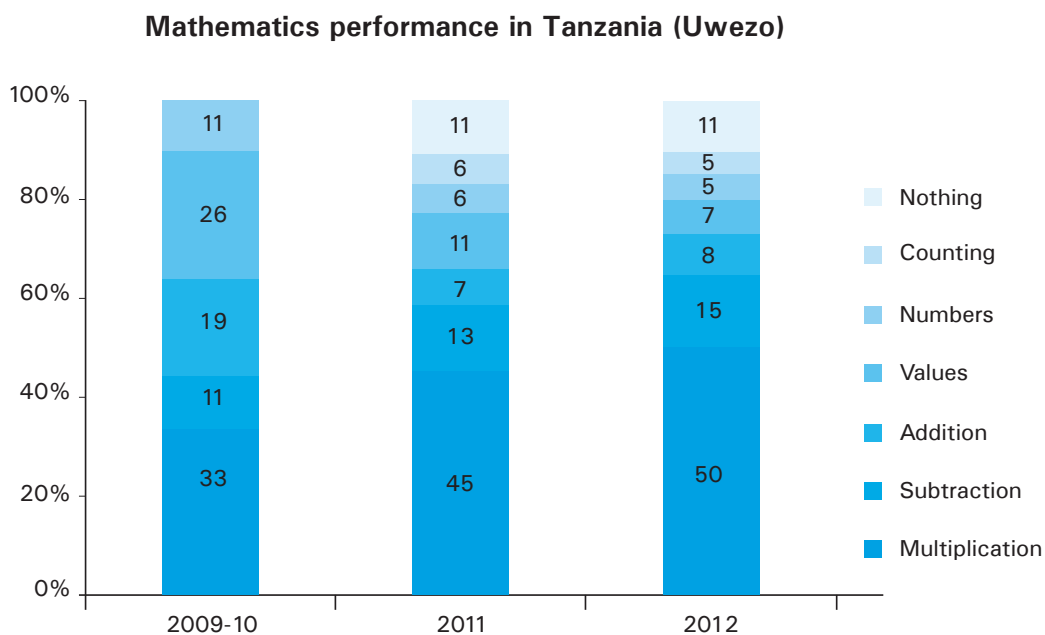
Figure 4. Trends in English performance across time for students in Kenya (Uwezo)



2.2.2 Uwezo numeracy trends

Similar to reading, seven common groups of tasks of increasing difficulty were also defined for mathematics, including 'counting', 'numbers', 'values', 'addition', 'subtraction', 'multiplication' and 'division'. An example for mathematics performance in Tanzania over time is shown below (see Figure 5). Little variation is evident across the three years in the percentage of students not able to complete anything (11 per cent). At the other end of the performance spectrum, there is an increase from approximately one-third of students able to complete multiplication in 2009–10 to approximately one-half of all students in 2012, a significant improvement that remained after age and gender of the populations were taken into account.

Figure 5. Trends in Mathematics performance across time for students in Tanzania (Uwezo)



For Kenya, we compared mathematics trend results between Uwezo 2011 and 2012 and found little difference in performance across the two years.⁶⁸ In Uganda, there was a drop from 32 per cent to 26 per cent between 2011 and 2013 in the percentage of students who were able to successfully complete the most difficult task, which appears to be related to the increased percentage of students achieving their highest task performance at lower levels.

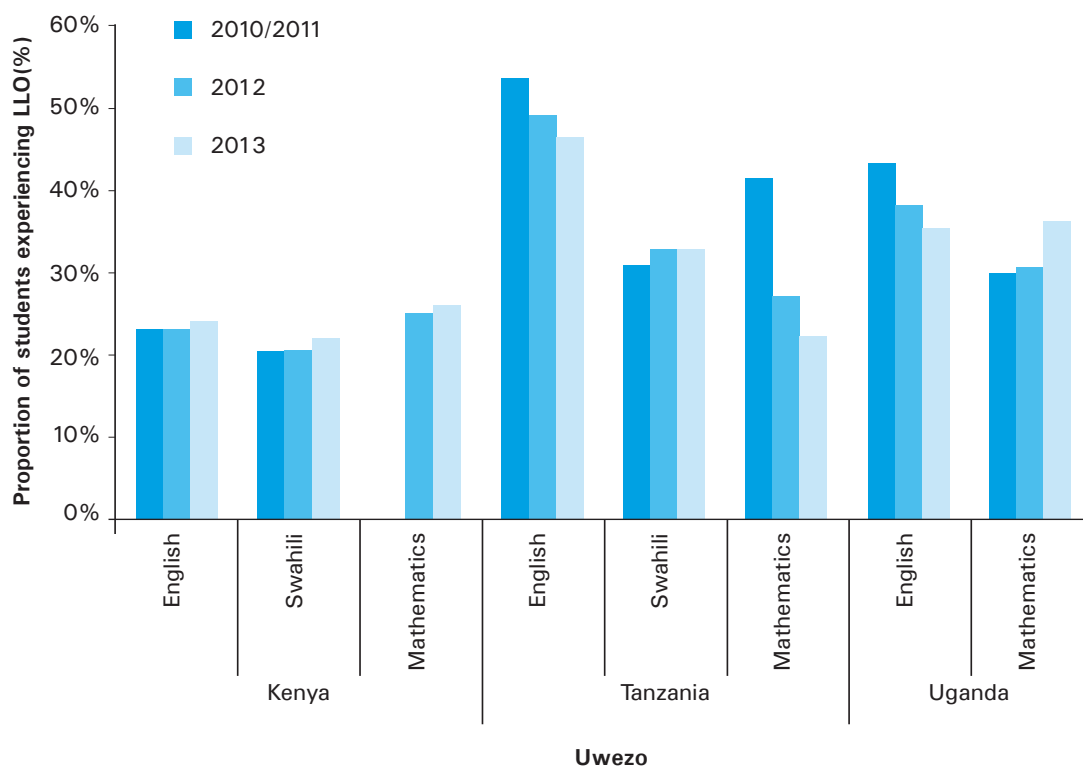
Tables with detailed results for the mathematics performance trends over time in Kenya and Uganda are presented in Appendix IV.

⁶⁸ The Uwezo division task was not administered in Kenya in 2009–10. The nature of the mathematics performance data from 2009–2010 is therefore not consistent with data from the other two assessments, and was not included in the trends analysis.

2.2.3 Uwezo: trends for children experiencing LLOs in literacy and numeracy

In addition to monitoring trends in performance at each task level, the trends in the percentage of children experiencing LLOs over time were analysed, using the same criteria reported in Chapter 2.1 and defined in Appendix II. This analysis allows us to assess whether any potential interventions or policy changes at national levels had the desired effects on children performing at the lowest levels. The trend data for the proportions of students who are experiencing LLOs for each domain, for each country participating in UWEZO appears below (see Figure 6).⁶⁹

Figure 6. Trends in proportions of students experiencing LLOs across Uwezo countries



We found little change over time in the relative percentages of students experiencing LLOs in Kenya for each of the three domains. A slight but significant increase can be seen from 2009 to 2010 in Swahili.

In Tanzania, however, a significant reduction in the percentage of students experiencing LLOs in English and mathematics is observed from 2009/2010 to 2011 and also from 2011 to 2012. Conversely, a greater percentage of children experienced LLOs in Swahili from 2009/2010 to 2011.

⁶⁹ Mathematics data for Kenya in 2010/2011 was not comparable with data for 2012 and 2013, and is thus not reported in this figure.

In Uganda, we observed a reduction in the proportions of children experiencing LLOs in English across the three time periods. For mathematics, however, we noted a significant increase from 2012 to 2013 in the proportions of children experiencing these problems.

Overall, when looking at the trend patterns from Uwezo across domains, there are instances where a higher percentage of students were able to complete the even more difficult tasks. On the other hand, there are instances where the proportion of students unable to complete the most basic of task changed. The reasons for these changes are not clear, nor whether they are positive or negative. To comprehend them, a better understanding is needed of the policy, financing and socio-political environments for education in each country – and how these developed over the years during which Uwezo was conducted.

2.2.4 Other literacy and numeracy trends for the ESA region

Additional information on trends in performance for the region is available from SACMEQ and PIRLS. Table 12 summarises the direction of major trend increases or decreases for the 15 entities that participated in SACMEQ reading and mathematics assessments for 2000 and 2007. A major increase or decrease is defined as greater than 10 scale points (Hungu et al., 2011). Lesotho, Mauritius, Namibia, Swaziland and Tanzania (Mainland) all had increases in both reading and mathematics of more than 10 points. In Botswana and Tanzania (Zanzibar), only reading scores improved by this margin over the time period, while in Malawi only mathematics scores improved to this extent. A decrease of more than 10 points was found in Mozambique for both reading and mathematics, and only for mathematics in Uganda (Hungu et al., 2011).

Grade 5 students from South Africa participated in the PIRLS study for two cycles (2006 and 2011). Although reading performance for students in 2011 was higher than for students in 2006, this difference was not significant for either English or Afrikaans (Howie et al., 2012). Girls outperformed boys in both cycles, but the gender difference fell from 37 scale points in 2006 to 26 scale points in 2011.

The limited data available for this aspect of learning outcomes make any generalisations difficult for the ESA region. Conclusions based on improvement or decline in student abilities should only be considered at the national level, with careful consideration given to national contextual factors. The contextual background differences of students across years should also be taken into account when interpreting the results.

Table 12. Direction of trends in SACMEQ reading and mathematics scale scores from 2000 to 2007

Country	Reading	Mathematics
Botswana	▲	○
Kenya	○	○
Lesotho	▲	▲
Malawi	○	▲
Mauritius	▲	▲
Mozambique	▼	▼
Namibia	▲	▲
Seychelles	○	○
South Africa	○	○
Swaziland	▲	▲
Tanzania (Mariland)	▲	▲
Tanzania (Zanzibar)	▲	○
Uganda	○	▼
Zambia	○	○
Zimbabwe	○	•

- ▲ Increase of more than 10 points from 2000 to 2007
- ▼ Decrease of more than 10 points from 200 to 2007
- Less than 10 point increases or decreases from 200 to 2007
- Assessment not administered in 2007

Source: Hungi et al., 2011



3. Improving learning outcomes in the ESA region: Effective country level practices

Low learning outcomes in literacy and numeracy – along with a considerable proportion of disadvantaged students not reaching basic skills in these domains – are among the many challenges for primary education in the ESA region. The focus of this chapter is on country-level practices that have proved to be effective in improving the literacy and numeracy learning outcomes of disadvantaged primary students in the ESA region.

When speaking about the ‘effectiveness’ of a programme to improve student performance, two main aspects need to be considered:

- The characteristics and strategies of the programme that were developed and implemented to improve student performance, such as teacher training in core reading skills, producing reading materials in the local language, introducing a reading buddy;
- The characteristics of the programme evaluation design that allows for the measurement of the impact of the programme on student performance, such as using randomised control trials, and establishing a baseline, mid-line and end-line of student performance using adequate performance measures.

Both aspects are important. While the former make a programme effective, the latter provides evidence of a programme’s effectiveness. The more closely that the evaluation design is aligned with the programme’s intended goals – and the higher the quality of the tools to measure performance – the more valid and meaningful the results, which can then feed into strengthening educational interventions to improve student learning outcomes.

There is a substantial body of literature for the ESA region on the impact of practices to increase the quantitative aspects of education quality – such as access, enrolment and retention rates. In contrast, there are few reports on programmes to improve student learning outcomes and its impact. This observation is not particular to our study. A review of 115 impact evaluation studies in 33 low- and middle-income countries (Murnane and Ganimian, 2014) found a variety of policies that were effective in increasing enrolment of students from low-income families (Murnane and Ganimian, 2014, p. 43). Strategies to improve the quality of education – and hence student performance – were found to be more complex and thus more challenging to undertake. The review revealed that findings about the impact of strategies on student performance are often inconsistent, and that ‘blanket statements about the effectiveness of particular reform strategies are neither accurate nor helpful’ (Murnane and Ganimian, 2014, p. 44). Also, interventions may have different effects on different groups targeted. A study in Kenya, for example, found that low- and high-achieving students derived very different benefits from free English textbooks (Glewwe, Kremer and Moulin, 2009). It is therefore important to consider the effects of an intervention for specific groups or sub-groups, in order to understand whether the same intervention would have a similar outcome with a different population (Murnane and Ganimian, 2014, p. 44).

We take these considerations into account in the following analysis of effective country-level practices that focus on the improvement of learning outcomes in the literacy and numeracy of

disadvantaged children in primary education in ESAR. We selected the programmes based on three main principles:

- The programme aims at improving learning outcomes in literacy and numeracy.
- The programme targets disadvantaged children in primary education.
- The impact of the programme on children’s literacy and numeracy learning outcomes has been evaluated.

3.1 Country-level programmes analysed

Altogether 10 programmes were identified, in seven out of the 21 ESA countries that effectively improved learning outcomes in the literacy and numeracy of disadvantaged children in primary education. An overview of these programmes is presented in Appendix V (see Table 23, Appendix V). It is important to note that the programmes presented are examples and do not claim to be exhaustive.⁷⁰

We categorised the programmes into four groups according to their main objectives. We did this to assist the analysis and help readers identify programmes that match their interests:

- Early grade literacy or numeracy programmes (7):
 - Literacy Boost in Ethiopia, Malawi and Mozambique
 - Reading to Learn in Kenya and Uganda
 - Primary Maths and Reading Initiative (PMRI) in Kenya
 - Malawi Teacher Professional Development Support (MTPDS)
- School improvement programme (1):
 - JET’s School Improvement Programme in South Africa
- Early childhood development programmes (2):
 - ECD component of the Early Literacy Project in Mozambique
 - Early Literacy and Maths Initiative (ELMI) as part of the Innovation for Education Programme in - Rwanda.

The majority of the identified programmes aim at improving early grade literacy or numeracy (or both). Disadvantaged children targeted in these programmes are low performers from a low socio-economic background or from economically disadvantaged or remote areas. Literacy Boost in Mozambique addresses children affected by HIV/Aids. Young children in communities affected by HIV/Aids are also the target group of the ECD programme in Mozambique. ELMI targeted preschool children, including children from remote areas without access to ECD programmes. JET’s school improvement programme was designed for schools in economically disadvantaged rural areas.

All programmes were evaluated using either Randomised Control Trial (RCT) or quasi-experimental designs. In these settings, students being studied are allocated (or ‘randomly allocated’ in the case of RCT) to the intervention under study and compared with control groups receiving no intervention. In all programme evaluations a baseline and end-line was conducted to measure the progress of performance. Employing such an experimental evaluation design, with pre- and post-measurement of performance, ensures that these features are built in from the very beginning of the programme design. This is commonly observed in the example programmes.

⁷⁰ A variety of programmes are implemented in the ESA region, as well as in countries where no examples have been identified for this analysis. In general, we observed a lack of evaluation reports or other documentation about the impact of a particular programme on learning outcomes; another 10 programmes were identified during this study where evaluation, including measurement of learning outcomes, was still in progress. They were not included.

3.2 Key strategies for success

One common feature among the programmes we investigated is the holistic approach they take, in which a number of interventions to improve learning outcomes of disadvantaged children are implemented at different levels. Viewing a programme holistically means seamlessly weaving together activities for teachers' professional development, provision of teaching/learning materials, community mobilisation, and capacity-building at the system level. The interventions flowing from such an approach would work in harmony at the system-, community- and school-levels. The one shortcoming is that holistic approaches make it difficult to assess which strategies are successful and which, in turn, pose challenges for scaling up an intervention. Another shortcoming is the difficulty of ensuring all stakeholders involved are equally committed to the intervention.

Common intervention strategies identified for the four programme groups are discussed in more detail in the following sections.

3.2.1 Early grade literacy and numeracy programmes

The main interventions we observed in the early grade literacy and numeracy programmes are teacher training on reading/mathematics instruction; provision of teaching/learning materials; production of reading materials in the local language; and community- and home-based reading activities that increase access to reading materials for children in and out of school.

All seven programmes reported significant improvement in the reading and mathematics skills of students who were targeted for these interventions.

A number of interventions contributed to the success of Save the Children's Literacy Boost in Ethiopia, Malawi and Mozambique. In Ethiopia, having a reading buddy turned Literacy Boost non-reading students into readers. In Malawi, Literacy Boost employed a three-pronged approach that included: '(1) use of assessments to identify gaps and measure improvements in the five core reading skills; (2) training of teachers in the national curriculum with an emphasis on core reading skills; and (3) community action through community mobilisation to support children's reading' (Save the Children, n.d., p. 2). The success of the programme is credited to the use of these three measures.

Reading to Learn was implemented in the official languages of reading instruction in the early primary grades, i.e., Lango in Uganda and Swahili in Kenya. At the heart of the Reading to Learn programme is a 'five-step scaffolding approach to literacy instruction, building from a conceptual understanding of stories, to decoding letter-sound relationships and eventually writing new sentences and stories' (Lucas, McEwan, Ngware and Oketch, 2014, p. 951). The coherence of this instructional model, which was aligned with materials, teacher training and well-targeted instructional interventions, helped ensure its success.

The Primary Maths and Reading Initiative (PRIMR) in Kenya contributed to the development and use of pedagogical materials and practices to improve children's foundational reading and mathematics skills. Materials developed were congruent with Kenyan curriculum documents and included detailed lesson plans based on best teaching practices identified through research,

student books, training manuals for Teacher Advisory Centre (TAC) tutors and teacher training videos, teacher-support mechanisms, student assessment tools and teacher observation tools.⁷¹ The end-line impact evaluation showed remarkable improvements in pupils' literacy and numeracy abilities, especially for pupils starting at the lowest levels. Learning outcomes in English and Kiswahili improved significantly; smaller improvements were seen in mathematics. Overall, girls performed at the same level as boys – if not better – especially in literacy.

Strategies found to be highly beneficial for the implementation of PRIMR were:

- consistent TAC tutors' visits to schools to support teachers (including its facilitation through reimbursement);
- regular professional development through brief trainings, with follow-up and refresher meetings;
- changing mindsets from traditional teaching to more active, student-focused approaches;
- accurate distribution of classroom materials to schools based on school enrolment data;
- planning and a sophisticated distribution network;
- limiting the number of schools that a TAC tutor is responsible for;
- provision of books at a 1:1 ratio;
- accommodating more literacy and numeracy instructional time during the week;
- keeping the transfer of teachers trained in PRIMR to a minimum (RTI, 2014).

The main purpose of the Malawi Teacher Professional Development Support (MTPDS) project was to enhance the instructional practices of teachers, especially for early-grade reading. The project focused mainly on supporting the lower primary sub-sector, with an emphasis on teacher skill development, classroom support, and materials development. The evaluation report showed that students who participated in the intervention achieved noticeable gains in performance, surpassing students in the control schools, with significant differences seen on all sub-tests. Two key strategies were identified as having the most significant impact:

- intensive coaching for teachers provided by the project's primary education advisors;
- continuity in support over time (Randolph, Nkhoma and Backman, 2013).

3.2.2 School improvement programmes

JET's School Improvement Programme aims at improving the efficiency of the educational system through a systemic approach to enhancing the functioning and educational performance of schools. 'The key assumption underlying the model is that educational outcomes will improve if teachers are effective and the teaching and learning environments are supported by effective school organisation, community involvement, and district support and monitoring' (JET Education Services, n.d., p. 9).

The school improvement model entails seven elements:

- 1) **Stakeholder mobilisation** in the community to support the improvement programme;
- 2) **Planning and organisation** to improve school management and the functioning of schools as organisations, including curriculum management, strategic planning and financial management;
- 3) **Teacher performance**, including awareness of teaching goals, focus on learning outcomes, access to efficient curriculum delivery systems and resources and provision of curriculum planning and delivery materials, school support visits and cluster-level activities;

71 Ben Piper and the PRIMR team, personal communication, 31 March 2015.

- 4) **Parent involvement** through a parent mobilisation programme, which includes setting up home study groups monitored by parents, and developing a practical guide on how parents should support their children's learning;
- 5) **District support** provided at two levels, the district office and the circuit involved in the project, to provide additional capacity for planning and programming of school support and monitoring activities, and for coordination with district-level activities;
- 6) **Teacher competence** (subject knowledge and teaching skills): monitoring, planning and facilitating of teachers' professional development;
- 7) **Research, monitoring and evaluation**, including ongoing monitoring conducted by the project schools and district officials.

The key assumption underlying the model is that educational outcomes will improve if teachers are effective, and if the teaching and learning environments are supported by effective school organisation, community involvement, and district support and monitoring (JET Education Services, n.d.). The impact evaluation shows significant improvement in learning outcomes. The mathematics and literacy performance of learners at project schools improved by five percentage points compared with those at non-project schools (JET Education Services, n.d.).

3.2.3 Early Childhood Development programmes

The main interventions observed in early childhood development programmes are teacher (and parent) training on the Early Childhood Development (ECD) approach and provision of teaching and learning materials. A distinctive feature of the Early Literacy and Maths Initiative (ELMI) is that it equips parents and caregivers with the tools to support their children in developing ELM-promoted skills that encourage playful activities. This 'ELM at Home' approach aims at extending opportunities to develop ELM skills at home, especially for those children with no access to ECD centres (Save the Children, 2014).

Children who benefited from both of these ECD programmes show significant improvement in their cognitive development. The ECD programme in Mozambique reports that preschool interventions in rural communities improved a number of important dimensions of child development. These included cognitive, fine motor and socio-emotional development, which contribute to higher levels of school readiness and significantly increased primary school enrolment at the appropriate age. The report concludes that low-cost, community-based preschool interventions, such as those studied in the Early Literacy Project, show potential for positively affecting early childhood development in rural African contexts (Martinez, Naudeau and Pereira, 2012).

For ELMI, the preliminary findings from the mid-line evaluation demonstrate the effectiveness of the ELMI programme on children's learning gains, regardless of whether the programme is implemented at home by parents or by teachers at an ECD centre. In investigations of the drivers of children's learning gains, typical background characteristics, such as maternal education level and socio-economic status, were found to play a role. More importantly, the time spent in play activities at home at mid-line showed the most consistent relationship with skill growth. The mid-line report concludes that the strong relationship between playful activities between parents and children and learning gains highlights the benefits of children engaging in developmentally appropriate play at early ages (Save the Children, 2014).



4. A macro theory of change

Based on the evidence collected for this study, we developed a macro theory of change, which is aimed at monitoring and improving the literacy and numeracy performance of primary students in the region. The theory combines the main findings of the stock-taking and comparative analysis of the assessments we reviewed, the main messages derived from the data analysis and related literature on characteristics of children with learning outcomes and trends in performance over time, and the experiences of effective country-level practices in the ESA region.

The theory identifies an evidence-based monitoring and intervention cycle, in which the interdependent sequence of assessment, analysis and action – known as the ‘three A’s – forms the basis of long-term and sustainable change in student performance (see Figure 7):

- **Assessments**, i.e., the systematic, strategic and regular collection of data on educational outcomes and factors related to these outcomes. Assessments form the basis of the evidence-based monitoring and intervention cycle.
- **Analysis** and interpretation of the findings on student performance, contexts, their relations and trends over time at the education policy level. This is to identify the factors to be addressed at the different levels of the education system in order to improve student performance.
- **Action**, the development of targeted educational interventions to improve progress for all learners, based on the factors and levels identified in the education policy analysis.

The three As are embedded in the conceptual framework of input, process and outcome factors at the different levels of an education system, which forms an integral component of the evidence-based monitoring and intervention cycle (see Figure 7).⁷²

For assessments, the conceptual framework includes the classification and definition of the main factors to be considered, and offers guidance on the levels at which data needs to be collected.

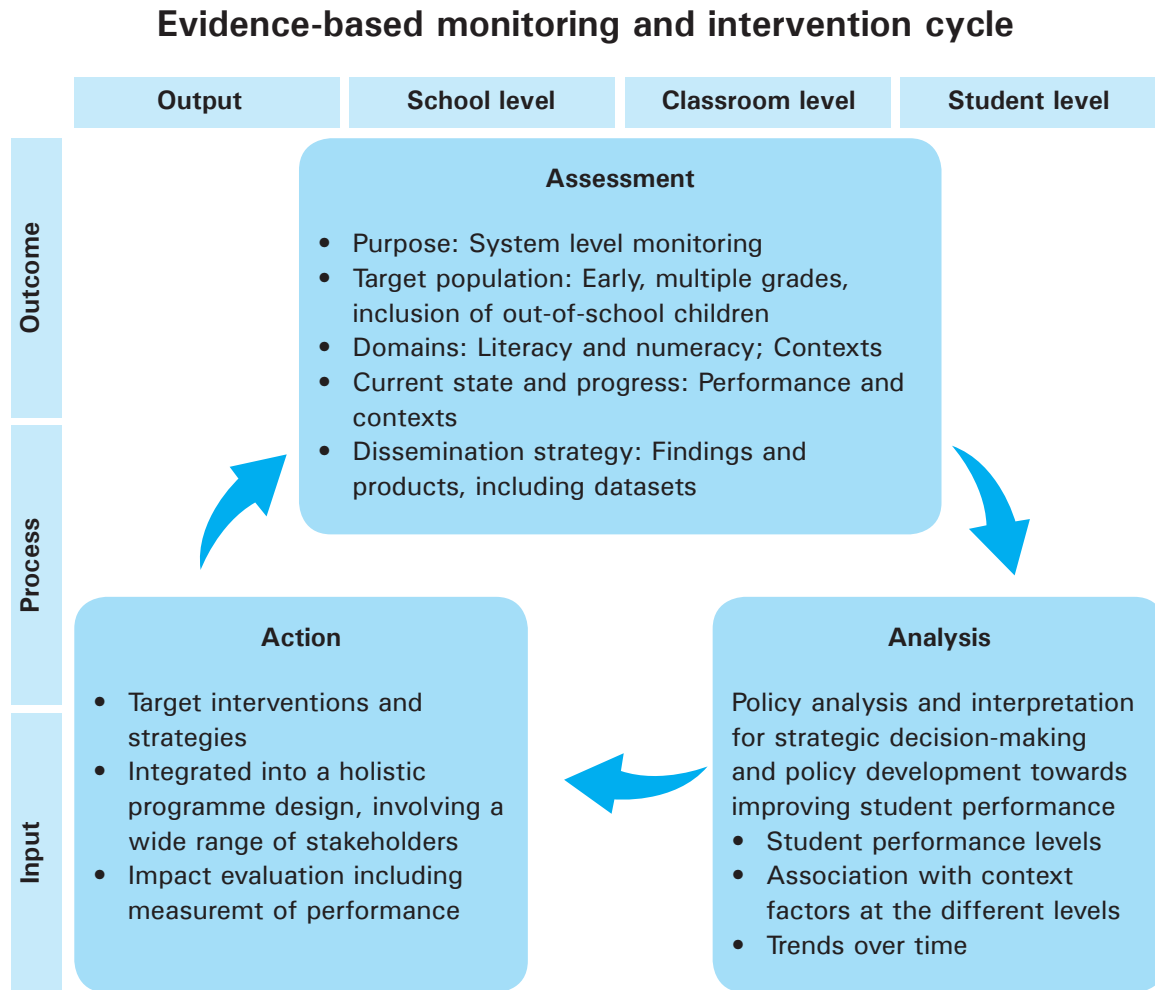
For analysis, the framework allows for the identification of factors to be addressed at the different levels of the education system in order to propel change, as well as their stability and suitability for change.

For action, the framework helps identify the programmes required and the level at which they should be implemented. For example, it helps policymakers decide which input, process and outcome factors need to be addressed in policy decisions at the system and school level, and which ones can best be shaped by school development activities at the school and classroom level.

Through the systematic, strategic and regular collection of data through assessments, progress in student performance is monitored, providing findings to be analysed and interpreted on policy level, which can then be transformed into action to further improve student performance.

⁷² The two-dimensional taxonomy is outlined in the conceptual framework in the introduction to this report.

Figure 7. A macro theory of change. An evidence-based monitoring and intervention cycle as premise for change: assessment, analysis, action



Key outcomes of the study informing the three A's

In our study, we found the following elements of the three A's important for the effectiveness of the monitoring and intervention cycle.

Assessment

Purpose

With the majority of the assessments conducted in the ESA region aimed at system-level monitoring (66 per cent), it is important that regular assessments be undertaken to monitor student performance levels and related context factors over time.

Target population

One of the initial policy decisions must be to define the target population(s) to be assessed. Grades 2 and 3 are the most frequently targeted in the ESA region. Nearly half of the assessments test multiple grades. Apart from Uwezo, which is a household-based assessment, all other assessments are school-based. The school setting excludes a significant proportion of children of primary school age who are out of school. Considerations on how to include out-of-school children in the assessment are necessary in order to gain a full picture of children's performance levels in a country or region.

Domains and contexts

Among basic competencies, literacy and numeracy are the most widely assessed. Two-thirds of the assessments in the ESA region assess both. Frameworks that clearly define and operationalize what educational themes are to be assessed are invaluable. They can help guide test development, data analysis and interpretation, and support a common understanding of the content and scope of the assessment among all stakeholders. In addition to performance data, it is essential to collect context data in order to explore the relationships between learning outcomes and background factors, and to identify factors relevant to change. The conceptual framework, which sets out relevant input, output and process factors on the different levels of the education system, can help ensure that the intervention instruments are well targeted in terms of scope and information source, e.g., students, teachers, principals, parents, EMIS.

Current state and progress: Performance and contexts

In order to draw conclusions about performance in literacy and numeracy and to identify which factors to address at which levels to initiate change, it is essential that competency levels and benchmarks be defined. In addition, to monitor progress over time, the relationships between context factors and performance must be understood.

In more than half of the assessments (55 per cent), results are presented with reference to competency levels or benchmarks. The description of the skills and knowledge required for the different competency levels provide a concrete understanding of what students can do. Based on this information, measures can be established to address specific learner needs at each level, for example, developing instructional strategies or identifying areas for professional training.

In order to accurately monitor progress over time, learning outcomes can only be compared between different implementations of the same assessment that have the same key features, such as assessment framework, design, target population and sampling design, and conditions of administration across multiple cycles. The contextual background differences of students across years in a study should also be taken into account when interpreting the results.

The same applies for comparisons across countries or regions. At present, the majority of the assessments in the ESA region are conducted at the national level, 29 per cent at the regional and 4 per cent at the international level.⁷³ An innovative approach to compare student performance across different assessments and contexts requires the development of common learning metrics for literacy and numeracy based on international and regional benchmarks for national achievement results and curricular expectations.

73 We view the purpose of the EGRA/EGMA implementations from the stock-taking as either system diagnostics or programme evaluation.

Item response theory (IRT) is an essential technique for scaling cognitive data in order to establish competency levels and effectively monitor performance on system level across contexts and over time. At present, IRT is not often used in the region. However two country case studies included in our study where IRT was used (ZELA in Zimbabwe and LARS in Rwanda) are good examples of innovative capacity-building programmes with a focus on data analysis techniques.

Dissemination

To engage stakeholders and instigate change, a clearly articulated dissemination strategy for the assessment findings is essential. Dissemination strategies were applied in the assessments that we reviewed. The majority of the assessments (71 per cent) made their results reports publicly available. Approximately half also released to the public results summaries, press communiqués and policy briefs. Some also made public fully documented datasets, which enable further independent analyses.

Analysis

The analysis and interpretation of assessment findings on student performance, context factors and trends over time can inform strategic decision-making and policy development. For example, the analysis we undertook of learning outcomes in the ESA region showed the proportion of students who experienced LLOs ranged from 18 per cent to 40 per cent for numeracy and 18 per cent to 50 per cent for reading literacy across datasets, countries and year levels. The message we take from these results is that individual student characteristics, home background and resourcing (at both the student and school level) are important factors that affect the ability of students to achieve basic levels of competency in literacy and numeracy. The results suggest that students would benefit from increased resourcing at the school level. At home, activities aimed at increasing student engagement with their studies would also have a profound impact on learning outcomes. The findings also suggest that encouraging parents to involve their children in literacy and numeracy activities before they start school, giving them more school-related responsibilities, and providing them with external tuition (if this was somehow made more readily available for those in need) would all positively impact their children's learning development.

Data that show improved learning outcomes in literacy and numeracy over time are indicative of systems that have helped improve overall student capabilities. Examining the education policies, financing and socio-political environments that may have contributed to these improvements would be helpful, as would studying the learning environments in countries where negative trends in outcomes have been observed.

Action

For our study, we examined 10 country-level programmes that have proven effective at improving literacy and numeracy learning outcomes of disadvantaged primary students.

Overall, common features included:

- Interventions and strategies targeted to a particular group of disadvantaged children, including children with low literacy achievement, children from disadvantaged socio-economic backgrounds, marginalised children in slums and non-formal settlements, and children in economically disadvantaged rural areas.

- A holistic programme design that addresses as many levels of the education system and stakeholders as possible. Such an approach typically involves teachers' professional development, provision of teaching/learning materials, community mobilisation and capacity-building at system level. This is a multi-level approach to programme implementation, which provides interventions at system-, school- and community-levels.
- Impact evaluation, including the measurement of progress in learning outcomes to provide evidence of success. Programmes with randomised control trials (RCT) or quasi-experimental designs require baseline, mid-line and end-line studies to ensure a built-in monitoring process from the beginning of the study.

Within programmes that share a particular purpose, we identified a number of key strategies that contributed to the success of country-level practices in the ESA region.

Effective early-grade literacy/numeracy programmes included teacher training on reading/mathematics instruction, aligned with the provision of teaching and learning materials and the production of reading materials in the local language, more active, student-focused approaches, the use of assessments, well targeted instructional interventions (e.g., students having a reading buddy to support their learning to read), increased instructional time, and community support for children's reading.

Additionally, programmes that aimed at a whole-school improvement were shown to have a significant impact on learning outcomes. Effective strategies included strong school organisation, community involvement, district support and monitoring and supportive learning environments.

Effective Early Childhood Development programmes comprised of teacher (and parent) training based on a particular model (e.g. encouraging playful activities involving early literacy and numeracy skills) and provision of teaching materials and tools also helped support early literacy and maths skill development.

These findings underline the importance of exposing children to a rich learning environment in their early years. Combined with an effective organisation that supports teaching and learning at all levels and equips teachers (and parents) with instructional approaches and educational material, these factors establish a sound basis for literacy and numeracy skills development.

Conclusions

Synthesising the main findings from this study, we developed a macro theory of change, anchored in the 'three A's approach (assessment, analysis and action) and aimed at initiating a long-term and sustainable change in student performance. While substantial research has been undertaken to identify the factors that contribute to student school attendance, more research is required to understand how student learning can be improved, particularly in developing countries. We need to deepen our knowledge of where students are at in their learning and how performance progresses over time so that effective targeted interventions can be developed. In order to do this, assessment programmes must be undertaken to provide quality comparable data across populations, between grades and over time. Finally, the results of the assessments must be integrated into education reform agendas, so that human and financial resources to address children's needs are efficiently deployed.



References

- ACER. (2014a). The Early Grade Reading Assessment: Assessing children's acquisition of basic literacy skills in developing countries. Assessment GEMS Series No. 2. ACER Press, Melbourne. <http://www.acer.edu.au/files/AssessGEMs_EGRA.pdf>
- ACER. (2014b). Inception report on specific methodology: Consultancy service for improving quality education and children's learning outcomes and effective practices in the Eastern and Southern Africa region. ACER Press, Melbourne.
- ACER. (2015). The Southern and Eastern Africa Consortium for Monitoring Educational Quality. Assessment GEMS Series No. 8. ACER Press, Melbourne.
- ACER. (n.d.). Learning assessments at a glance. ACER Press, Melbourne. Viewed 30 March 2015, <<http://www.acer.edu.au/gem-la>>.
- ACER, and ZIMSEC. (2013). Zimbabwe Early Learning Assessment (ETF programme) Base-line Study.
- ACER, and ZIMSEC. (2015). Evaluation of the Education Development Fund Programme: Zimbabwe Early Learning Assessment (ZELA). 2014 Monitoring Report.
- American Institutes for Research (AIR). (2012). Ethiopia Early Grade Reading Assessment - Data Analytic Report: Language and Early Learning: American Institutes for Research (AIR).
- American Institutes for Research (AIR). (n.d.). Namibian National Standardized Achievement Test (NSAT). Retrieved 17 March, 2015, from <http://www.air.org/project/namibian-national-standardized-achievement-test-nsat>
- Beattie, K. (2014). Early Grade Reading Assessment results Somalia Programme: Concern Worldwide.
- Beattie, K., & Grogan, R. (2013). Results of the Early Grade Reading Assessment conducted in five schools in Mogadishu schools: Concern Worldwide.
- Biggs, J. (1999). Teaching for Quality Learning at University. Philadelphia, PA: Society for Research into Higher Education and Open University Press.
- Biggs, J., and Moore, P. (1993). The Process of Learning (3rd edn). Australia: Prentice Hall.
- Brombacher, A., Nordstrum, L., Davidson, M., Batchelder, K., Cummiskey, C., & King, S. (2014). National Baseline Assessment for the 3Rs (Reading, Writing, and Arithmetic) Using EGRA, EGMA, and SSME in Tanzania: Study Report.
- Cao, Y., Dowd, A. J., Mohammed, O., Hassen, S., Hordofa, T., Diyana, F., & Ochoa, C. (2011). Literacy Boost, Dendi Ethiopia: Three month report: Save the Children.
- Collins, P., Galbert, P. D., Hartwell, A., Kochetkova, E., Mulcahy-Dunn, A., Nimbalkar, A., and Ralaingita, W. (2012). Pupil Performance, Pedagogic Practice, and School Management: An SSME Pilot in Zambia: RTI, North Carolina.
- CONFEMEN. (2010a). Rapport PASEC–Burundi 2010: Enseignement primaire: Quels défis pour une éducation de qualité en 2015? CONFEMEN, Dakar, Senegal.

- CONFEMEN. (2010b). Rapport PASEC–Union des Comores: Diagnostique et préconisations pour une scolarisation universelle de qualité. CONFEMEN, Dakar, Senegal.
- CONFEMEN. (2010c). RAPPORT PASEC Union des Comores 2010. Diagnostic et préconisations pour une scolarisation universelle de qualité. PASEC/CONFEMEN, Comoros.
- CONFEMEN. (n.d.). Programme d'Analyse des Systèmes d'Éducatif de la CONFEMEN (PASEC). Viewed 16 March 2015, <<http://www.confemen.org/le-pasec/>>
- Constitution of Zimbabwe Amendment (No. 20), 2013. (2013).
- Creemers, B., and Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice, and Theory in Contemporary Schools*. Routledge, London.
- Crouch, L., Korda, M., & Mumo, D. (2009). *Improvements in Reading Skills in Kenya: An Experiment in the Malindi District*: RTI.
- Department of Basic Education Republic of South Africa. (2011). *Report on the Annual National Assessments of 2011*.
- Department of Basic Education Republic of South Africa. (2012). *Report on the Annual National Assessment of 2012: Grades 1-6 & 9*.
- Department of Basic Education Republic of South Africa. (2013). *Report on the Annual National Assessment of 2013: Grades 1-6 & 9*.
- Department of Basic Education Republic of South Africa. (2014). *Report on the Annual National Assessment of 2014: Grades 1–6 and 9*. Department of Basic Education, Pretoria.
- DeStefano, J., Ralaingita, W., Costello, M., Sax, A., & Frank, A. (2012). *Task Order 7: Early Grade Reading and Mathematics in Rwanda*: RTI.
- Dowd, A. J., & Fonseca, J. (2012). *Literacy Boost Results Mozambique*: Save the Children.
- Dowd, A. J., and Mabeti, F. (2011). *Literacy Boost Results Malawi: Year 2 report*. Save the Children, Washington DC.
- Dowd, A. J., Wiener, K., & Mabeti, F. (2010). *Literacy Boost Results Malawi: Year 1 report*: Save the Children.
- Examinations Council of Lesotho (ECOL). (n.d.). *Examinations Council of Lesotho: Research and Publications*. Retrieved 17 March, 2015, from <http://www.examsCouncil.org.ls/research.aspx>
- Examinations Council of Zambia. (2015). *Zambia - National Assessment Survey*.
- fhi 360 Education Policy and Data Centre. (2015). *EPDC Policy Brief: Mapping national assessments*. Education Policy and Data Centre, Washington DC.
- fhi 360 Education Policy and Data Centre. (n.d.-a). *Data*. Retrieved 2 April, 2015, from <http://www.epdc.org/data>
- fhi 360 Education Policy and Data Centre. (n.d.-b). *National Education Profiles*. Viewed 2 April 2015, <<http://www.epdc.org/tags/national-education-profiles>> .
- Friedlander, E., Candiru, S., & Dowd, A. J. (2010). *Literacy Boost Uganda: Baseline Report 2010*: Save the Children.

- Friedlander, E., Hordofa, T., Diyana, F., Hassen, S., Mohammed, O., and Dowd, A. J. (2012). Literacy Boost, Dendi Ethiopia: Endline II. Save the Children, Washington DC.
- Fullan, M. (2006). Change theory: A force for school improvement. Seminar Series Paper no. 157, November. CSE, Melbourne.
- Garrouste, C. (2011). 'Explaining learning gaps in Namibia: The role of language proficiency.' *International Journal of Educational Development*, 31(3), 223–233. doi: 10.1016/j.ijedudev.2010.06.016
- Glewwe, P., Kremer, M., and Moulin, S. (2009). 'Many Children Left Behind? Textbooks and Test Scores in Kenya.' *American Economic Journal: Applied Economics*, 1(1), 112–135.
- Global Partnership for Education. (2012). Results for Learning Report 2012: Fostering Evidence-Based Dialogue to Monitor Access and Quality in Education. GPE, Washington DC.
- Gove, A., and Wetterberg, A. (2011). The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy. RTI International, North Carolina.
- Government of Zimbabwe. (2009). Short Term Emergency Recovery Programme (STERP): Getting Zimbabwe moving again. Government of Zimbabwe, Harare.
- Guajardo, J., Night, C. S., Heijnen, E., Komakech, C., P'kech, E. A., Odwong, O. D., . . . Dowd, A. J. (2010). Literacy Boost Uganda: Midline Report 2012: Save the Children.
- Hassen, S., & Friedlander, E. (2012). Literacy Boost Results Ethiopia: Save the Children.
- Hivos/Twaweza. (2014). Are Our Children Learning? Literacy and Numeracy across East Africa 2013. Uwezo East Africa at Twaweza, Kenya.
- Howie, S., Staden, S. v., Tshele, M., Dowse, C., and Zimmerman, L. (2012). PIRLS 2011. South African Children's Reading Literacy Achievement Report. Centre for Evaluation and Assessment, University of Pretoria, Pretoria.
- Human Sciences Research Council South Africa. (n.d.). Testing, testing First national assessment of Grade 9 pupils shows much work lies ahead. Retrieved 20 March, 2015, from <http://www.hsrc.ac.za/en/review/september-2011/testing-testing#sthash.q42e9GXL.dpuf>
- Hungi, N. (2011). Characteristics of school heads and their schools. UNESCO, Paris.
- Hungi, N. (2011a). Accounting for variations in the quality of primary school education. Retrieved from: http://www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iii/working-papers/07_multivariate_final.pdf
- Hungi, N. (2011m). Characteristics of grade 6 pupils, their homes and learning environments SACMEQ Working Paper. Paris: SACMEQ.
- Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., van Cappelle, F., . . . Vellien, J. (2010). SACMEQ III project results: Pupil achievement levels in reading and mathematics. SACMEQ, Paris.
- Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., van Cappelle, F., . . . Vellien, J. (2011). SACMEQ III project results: Levels and trends in school resources among SACMEQ school systems. SACMEQ Working Document. Paris: SACMEQ.

- Hungi, N., Ngware, M., and Abuya, B. (2014). 'Examining the impact of age on literacy achievement among grade 6 primary school pupils in Kenya.' *International Journal of Educational Development*, 39, 247–259.
- IEA. (2013a). PIRLS 2011. IEA, Amsterdam
- IEA. (2013b). PIRLS 2011 Contextual Questionnaires. Viewed 17 April 2015, <<http://timssandpirls.bc.edu/pirls2011/international-contextual-q.html>>.
- IEA. (2013c). TIMSS 2011. Viewed 7 August 2015, <<http://timssandpirls.bc.edu/timss2011/schedule.html>>
- IEA. (2013d). TIMSS 2011 Contextual Questionnaires. Viewed 17 April 2015, <<http://timssandpirls.bc.edu/timss2011/international-contextual-q.html>>
- IEA. (n.d.). IEA Study Data Repository. Viewed 17 March 2015, <<http://rms.iea-dpc.org/>>.
- JET Education Services. (n.d.). Sustainable school improvement. JET, Zambia.
- Joncas, M. (2011). *Methods and procedures: PIRLS 2011 Target populations: IMSS & PIRLS* International Study Centre, Lynch School of Education, Boston College.
- Kaplan, R. M., and Saccuzo, D. P. (1997). *Psychological testing: Principles, applications and issues*. Brooks Cole Publishing Company, Pacific Grove, California.
- Kyriakides, L., and Creemers, B. (2006). 'Using the dynamic model of educational effectiveness to introduce a policy promoting the provision of equal opportunities to students of different social groups.' In D. M. McInerney, M. Dowson, and S. V. Etten (eds), *Effective Schools*, vol. 6, pp. 17–41. Information Age Publishing, North Carolina.
- Lesotho National Assessment of Educational Progress (LNAEP) Survey Report, 2010. (n.d.).
- Lucas, A. M., McEwan, P. J., Ngware, M., and Oketch, M. (2014). *Improving early-grade literacy in East Africa: Experimental evidence from Kenya and Uganda*. University of Delaware, Delaware.
- Majgaard, K., and Mingat, A. (2012). *Education in sub-Saharan Africa: A comparative analysis*. World Bank, Washington DC.
- Makuwa, D. (2005). *The SACMEQ II Project in Namibia: A Study of the Conditions of Schooling and the Quality of Education*. Ministry of Basic Education, Sport and Culture, Namibia.
- Makuwa, D. (2011). *Characteristics of grade 6 teachers SACMEQ Working Paper*. Paris: SACMEQ.
- Martin, M. O., and Mullis, I. V. S., (eds). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. TIMSS and PIRLS International Study Center, Boston College, MA.
- Martin, M. O., Mullis, I. V. S., Foy, P., and Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. TIMSS and PIRLS International Study Center, Boston College, MA.
- Martin, M. O., Mullis, I. V. S., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 International Results in Science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martinez, S., Naudeau, S., and Pereira, V. (2012). *The Promise of Preschool in Africa: A Randomized Impact Evaluation of Early Childhood Development in Rural Mozambique*. 31e, Washington DC.

- Mason, M. (2014). *Complexity Theory in Education Governance: Initiating and sustaining systemic change*. UNESCO Bureau of Education, Paris.
- Miksic, E., & Harvey, S. (2012). *Malawi National Early Grade Reading Assessment Survey: Midterm Assessment: RTI*.
- Ministère de l'Éducation Nationale et de la Recherche, & PASEC-CONFEMEN. (2010). *Rapport PASEC - Union des Comores: Diagnostique et préconisations pour une scolarisation universelle de qualité*. Dakar, Senegal: CONFEMEN.
- Ministère de l'Enseignement de Base et Secondaire, de l'Enseignement de Métiers, de la Formation Professionnelle et de l'Alphabétisation, & PASEC-CONFEMEN. (2010). *Rapport PASEC - Burundi 2010: Enseignement primaire: Quels défis pour une éducation de qualité en 2015?* Dakar, Senegal: CONFEMEN.
- Ministry of Education, Science, and Technology of Malawi. (2011). *The SACMEQ III project in Malawi*. Malawi Ministry of Education, Science and Technology, Malawi.
- Ministry of Education, Science and Technology of Malawi. (2014). *Monitoring Learning Achievement in Primary Education Malawi Report*. Malawi Ministry of Education, Science and Technology, Malawi.
- Ministry of Education of Angola, World Bank, A. R., Human Development Department, and Russia Education Aid for Development Programme. (2011). *Early Grade Literacy in Angola*.
- Ministry of Education of Ethiopia (FDRE). (2008). *Ethiopian Third National Learning Assessment of Eighth Grade Students*. Ministry of Education of Ethiopia, Addis Ababa.
- Ministry of Education of Ethiopia (FDRE). (2013). *Ethiopian 4th National Learning Assessment of Grades 4 and 8 Pupils: Data Analytic Report*. Addis Ababa, Ethiopia.
- MOESAC. (2006). *Education Amendment Act, 2005 (Chapter 25:04), H.B.6A, 2005*. Government of Zimbabwe, Harare.
- MOESAC. (2009). *Education at a Glance 2009*. MOESAC, Harare.
- MOESAC. (2011). *Education Medium Term Plan 2011–2015*. MOESAC, Harare.
- Moloi, M. (n.d.). *Mathematics achievement in South Africa: A comparison of the official curriculum with pupil performance in the SACMEQ II Project*.
- Moloi, M. Q., and Chetty, M. (2010). *The SACMEQ III project in South Africa: A study of the conditions of schooling and the quality of education: South Africa Country Report*. Department of Basic Education, Pretoria.
- Monyaku, B., & Mmereki, O. A. (2011). *The SACMEQ III project in Botswana: A study of the conditions of schooling and the quality of education: Botswana Ministry of Education and Skills Development, Division of Planning Statistics and Research*.
- Monyaku, B. (2012). *The SACMEQ III project in Botswana: A study of the conditions of schooling and the quality of education*. Botswana Ministry of Education and Skills Development, Botswana.

- Mullis, I., Martin, M., Foy, P., and Arora, A. (Eds.). (2012). *TIMSS 2011 International Results in Mathematics*. TIMSS and PIRLS International Study Center, Boston College, and International Association for the Evaluation of Educational Achievement (IEA). Chestnut Hill, MA.
- Mullis, I., Martin, M., Foy, P., and Drucker, K. (2012). *PIRLS 2011 International Results in Reading*. TIMSS and PIRLS International Study Center, Boston College, and International Association for the Evaluation of Educational Achievement (IEA). Chestnut Hill, MA.
- Mullis, I. V. S., and Martin, M. O. (Eds.). (2013). *PIRLS 2016 Assessment Framework*. TIMSS and PIRLS International Study Center, Boston College, and International Association for the Evaluation of Educational Achievement (IEA). Chestnut Hill, MA.
- Mullis, I. V. S., Martin, M. O., Foy, P., & (with Olson, J. F., Preuschoff, C., Erberber, E., & Galia, J.). (2008a). *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P., & (with Olson, J. F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008b). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 International Results in Reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy P. (2007). *IEA's Progress in International Reading Literacy Study in Primary School in 40 Countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A., Trong, K., and Sainsbury, M. (2009). *PIRLS 2011 Assessment Framework*. TIMSS and PIRLS International Study Center and International Association for the Evaluation of Educational Achievement. IEA, Amsterdam.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., and Preuschoff, C. (2009). *TIMSS 2011 Assessment Frameworks*. TIMSS and PIRLS International Study Center and International Association for the Evaluation of Educational Achievement. IEA, Amsterdam.
- Mungoi, D., Mandlante, N., Nhatuve, I., Mahangue, D., Fonseca, J., and Dowd, A. J. (2010). *Endline Report of Early Literacy among Pre-school and Primary School Children in Gaza Province*. Save the Children, Mozambique
- Murnane, R. J., and Ganimian, A. (2014). *Improving educational outcomes in developing countries: Lessons from rigorous evaluations: NBER Working Paper 20284*. National Bureau of Economic Research, Cambridge, MA.
- National Educational Assessments and Examinations Association of Ethiopia. (n.d.). *National Educational Assessments and Examinations Association*. Retrieved 17 March, 2015, from <http://www.nae.gov.et/HOME.aspx>
- Nhongo, K. (2014, 19 March). *Grade 5 NSAT results depressing*. Windhoek Observer. Retrieved from <http://observer24.com.na/national/3131-grade-5-nsat-results-depressing>
- Nyanguru, A., and Peil, M. (1991). 'Zimbabwe since Independence: A people's assessment.' *African Affairs*, 90(361), 607-620.

- OECD. (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing.
- Piper, B. (2010). *Uganda Early Grade Reading Assessment Findings Report: Literacy Acquisition and Mother Tongue*: RTI.
- Piper, B., & Mugenda, A. (2013). *The Primary Math and Reading (PRIMR) Initiative: Midterm Impact Evaluation*. North Carolina, USA: RTI.
- Pouzevara, S., Costello, M., & Banda, O. (2013). *Malawi National Early Grade Reading Assessment Survey: Final Assessment - November 2012*: RTI.
- Purves, A. C. (1987). 'The Evolution of the IEA: A Memoir'. *Comparative Education Review*, 31(1), 10–28.
- Randolph, E., Nkhoma, M., and Backman, S. (2013). *ABE/LINK Malawi Teacher Professional Development Support: Project M&E report*. USAID.
- Raupp, M., Newman, B., & Revés, L. (2013). *Impact evaluation for the USAID/Aprender la ler project in Mozambique: Baseline report: International Business & Technical Consultants, Inc. (IBTCI)*.
- Report on Monitoring Learning Achievements (MLA) in Grade 4 in Puntland and Somaliland. (2012).
- Results for Development. (n.d.). Center for Education Innovations. Viewed 2 April 2015, <<http://www.educationinnovations.org/>> .
- Ross, K. (2009). The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ). Paper presented at the Global Partnership for Education's 'Biennial Partnership Meeting' Copenhagen, April 20–21.
- RTI. (2004). *Eddata II: Education data for decision making: School management*. Viewed 17 April 2015, <<https://www.eddataglobal.org/management/index.cfm>> .
- RTI. (2010). *Ethiopia Early Grade Reading Assessment–Data Analytic Report: Language and Early Learning*. RTI International, North Carolina.
- RTI. (2011). *Malawi National Early Grade Reading Assessment Survey: Baseline Assessment*: RTI.
- RTI. (2012). *The Primary Math and Reading (PRIMR) Initiative: Baseline Report*. North Carolina, USA.
- RTI. (2014a). *EGRA Tracker*.
- RTI. (2014). *USAID/Kenya Primary Math and Reading (PRIMR) Initiative: Final Report*. RTI, North Carolina.
- RTI. (2015a). *EGRA Tracker*. In *EGRA_tracker_for_web_16June2015.xlsx* (Ed.).
- RTI (2015c, January 9). [Personal Communication, Request for data access].
- RTI. (n.d.). *EdData: Education data for decisionmaking*. Retrieved 17 March, 2015, from <https://www.eddataglobal.org/index.cfm>
- Rwanda Education Board. (2012). *Learning Achievement in Rwandan Schools (LARS)*. Rwanda Education Board (REB), Kigali, Rwanda.

- Rwanda Education Ministry. (2014). 2013 Education Statistical Yearbook. Ministry of Education, Kigali.
- SACMEQ. (2013). SACMEQ. Viewed 31 March 2015, <<http://www.sacmeq.org/>> .
- Sakala, C. T., & Chilala, M. M. (2007). The Role of the Zambia National Assessment Programme in Evaluating the Attainment of Educational Goals. Paper presented at the 33rd International Association for Educational Assessment Annual Conference, Baku, Azerbaijan. http://www.iaea.info/documents/paper_1162d22700.pdf
- Sasman, C. (2011, 19 July). Grade 7 tests show disappointing results. The Namibian. Retrieved from http://www.namibian.com.na/index.php?archive_id=82769&page_type=archive_story_detail&page=1580
- Satio, M. (2011). Trends in the magnitude and direction of gender differences in learning outcomes, SACMEQ Working Paper No .4. SACMEQ.
- Save the Children. (2014). Early Literacy and Maths Initiative (ELMI): Rwanda Midline Report. Save the Children, Washington DC.
- Save the Children. (n.d.). Literacy Boost. Viewed 17 March 2015, <<http://resourcecentre.savethechildren.se/>> .
- The Australian Council for Educational Research, & Zimbabwe School Examination Council. (2013a). Evaluation of the Education Development Fund Program - Zimbabwe Early Learning Assessment (ZELA): 2013 Monitoring Report.
- The Australian Council for Educational Research, & Zimbabwe School Examination Council. (2013c). Zimbabwe Early Learning Assessment (ETF program) Base-line Study.
- The Australian Council for Educational Research, & Zimbabwe School Examination Council. (2015). Evaluation of the Education Development Fund Program - Zimbabwe Early Learning Assessment (ZELA): 2014 Monitoring Report.
- The Kenya National Examinations Council. (n.d.). The Kenya National Examinations Council: NASMLA. Retrieved 17 March, 2015, from http://www.knec.ac.ke/main/index.php?option=com_phocadownload&view=section&id=43:nasmla
- The National Assessment Centre. (2010). Monitoring of learner achievement for class 3 in literacy and numeracy in Kenya: Summary of results and recommendations. The Kenya National Examinations Council, Nairobi, Kenya.
- The National Assessment Centre. (2010a). Monitoring of learner achievement for class 3 in literacy and numeracy in Kenya. Nairobi, Kenya: The Kenya National Examinations Council.
- The National Assessment Centre. (2010c). Monitoring of learner achievement for class 3 in literacy and numeracy in Kenya: Summary of results and recommendations. Nairobi, Kenya: The Kenya National Examinations Council.
- TIMSS and PIRLS International Study Center. (n.d.). TIMSS and PIRLS. Viewed 16 March 2015, <<http://timssandpirls.bc.edu/>> .
- Twaweza. (n.d.). Uwezo. Viewed 17 March 2015, <<http://www.uwezo.net/>> .
- Uganda National Examinations Board. (2010). The Achievement of Primary School Pupils in Uganda in Numeracy, Literacy in English and Local Languages. Uganda National Examinations Board, Kampala.

- Uganda National Examinations Board. (n.d.). Uganda National Examinations Board - Home. Retrieved 18 March, 2015, from <http://www.uneb.ac.ug/index.php?link=Home>
- UNESCO. (2008). EFA global monitoring report 2008—Education for all by 2015: Will we make it? UNESCO, Paris.
- UNESCO. (2010). EFA Global Monitoring Report: Reaching the marginalized: Regional overview: sub-Saharan Africa. UNESCO, Paris.
- UNESCO. (2014). Education for All in sub-Saharan Africa Assessment Report 2014. UNESCO Dakar Regional Office, Dakar.
- UNESCO. (2015a). EFA Global Monitoring Report 2015. EFA 2000–2015: Achievements and Challenges. UNESCO, Paris.
- UNESCO. (2015b). National Assessments [Excel file for Annex of 2015 EFA GMR report—in publication].
- UNESCO. (n.d.-a). Education For All Global Monitoring Report. Viewed 2 April 2015, <<http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/efareport/>> .
- UNESCO. (n.d.-b). International Bureau of Education. Viewed 2 April 2015, <<http://www.ibe.unesco.org/en.html>> .
- UNICEF Eritrea. (n.d.). Monitoring Learning Achievement in Eritrea.
- UNICEF Mozambique Country Office (2015). [Comments on the Improving Quality Education and Children’s Learning Outcomes and Effective Practices in the Eastern and Southern Africa region: Final Report].
- UNICEF. (2008a). News note: Zimbabwe education system in a state of emergency. Viewed 1 April 2015, <http://www.unicef.org/media/media_45950.html> .
- UNICEF. (2008b). Progress report for UNICEF’s education in emergencies and post-crisis transition programme. UNICEF.
- UNICEF Rwanda Country Office (2015). [Comments on the Improving Quality Education and Children’s Learning Outcomes and Effective Practices in the Eastern and Southern Africa region: Final Report].
- UNICEF Somalia. (n.d.). Report on Monitoring Learning Achievements (MLA) in Grade 4 in Puntland and Somaliland.
- UNICEF. (2011). The Education Transition Fund II, 2012–2015: Programme Document. UNICEF, Harare.
- UNICEF Zambia Country Office (2015). [Comments on the Improving Quality Education and Children’s Learning Outcomes and Effective Practices in the Eastern and Southern Africa region: Final Report].
- UNICEF. (2012). Zimbabwe 2012: Millennium Development Goals Progress Report. UNICEF Zimbabwe, Harare.
- UNICEF. (2013a). Country Office Portal: Annual report 2013 for Rwanda. ESARO.
- UNICEF. (2013b). Country Office Portal: Annual report 2013 for Zimbabwe.
- UNICEF. (2014). The Education Development Fund: Stronger systems, better outcomes: Sixth Progress Report. Harare: UNICEF Zimbabwe.

- UNICEF Eritrea. (n.d.). Monitoring Learning Achievement in Eritrea.
- UNICEF Zambia Country Office (2015). Comments on the Improving Quality Education and Children's Learning Outcomes and Effective Practices in the Eastern and Southern Africa region: Final Report. UNICEF.
- USAID. (2011). USAID Funded Malawi Teacher Professional Development Support (MTPDS) Program - 2010 Early Grade Assessment (EGMA): National Baseline Report 2010.
- Uwezo-Kenya. (2013). Are Our Children Learning? Annual Learning Assessment Report 2012. Nairobi, Kenya: Uwezo, WERK.
- Uwezo-Kenya. (2013). Survey Booklet 2013. Uwezo.
- Uwezo-Tanzania. (2013). Are Our Children Learning? Annual Learning Assessment Report 2012. Dar es Salaam, Tanzania: Uwezo, TEN/MET.
- Uwezo-Uganda. (2013). Are Our Children Learning? Annual Learning Assessment Report 2012. Kampala, Uganda: UWEZO Uganda.
- Uwezo. (2014). Are Our Children Learning? Literacy and Numeracy across East Africa 2013. Nairobi, Kenya: Uwezo, HIVOS/Twaweza.
- Uwezo. (2014). Datasets viewed November 19 2014, <<http://www.uwezo.net/>> .
- Vogel, I. (2012). Review of the use of 'Theory of Change' in international development: Review report. DFID, London.
- Wasanga, P. M., A. Ogle, M., and Wambua, R. M. (2012). The SACMEQ III project in Kenya: A study of the conditions of schooling and the quality of education. The Kenya National Examinations Council, Nairobi.
- Wasanga, P. M., Ogle, M. A., and Wambua, R. M. (2010). The Report on Monitoring Learner Achievement Study for Class 3 in Literacy and Numeracy. The Kenya National Examination Council, Nairobi.
- Wasanga, P. M., Ogle, M. A., & Wambua, R. M. (2012). The SACMEQ III project in Kenya: A study of the conditions of schooling and the quality of education: The Kenya National Examinations Council.
- World Bank. (2015). Zimbabwe. <<http://data.worldbank.org/country/zimbabwe>> .





Australian Council for Educational Research

