Ieva Raudonytė and Tuamanaia Foimapafisi

unesco
International Institute for Educational Planning

Using Learning Assessment Data for Educational Planning in Sub-Saharan Africa: A comparative analysis

Use of Learning Assessment Data

Ieva Raudonytė and Tuamanaia Foimapafisi

Using Learning Assessment Data for Educational Planning in Sub-Saharan Africa: A comparative analysis

Use of Learning Assessment Data

Paris 2022 UNESCO International Institute for Educational Planning

.....

Published by:
International Institute for Educational Planning
7–9 rue Eugène Delacroix, 75116 Paris
info@iiep.unesco.org
www.iiep.unesco.org
ISBN: 978-92-803-1453-3
Cover photo: Shutterstock
Typesetting: Regent Typesetting

unesco
International Institute for Educational Planning

Acknowledgements

This book was developed in the framework of the IIEP-UNESCO research programme on the use of learning assessment data in the education planning cycle, under the overall guidance and supervision of Hugues Moussy, a former head of the Research and Development team at IIEP-UNESCO. We would like to thank the researchers who contributed to implementing this study: Might Abreh, Djénabou Baldé, Omar Ceesay, Mariatou Faal, Fatimata Hydara, Gift Masaiti, Innocent Mulenga, Hertha Pomuti, Charmaine Villet, Ousmane Senghor, Modou Touray, and Winnie Yondo. This work could not have been done without the support of the ministries of education that enabled and facilitated the research work in their countries. We sincerely thank ministry officials and international partners for taking the time to participate in our interviews and/or focus group discussions. This document benefited greatly from the review and in-depth comments of Chloé Chimier, Michaela Martin, Mioko Saito, Lynne Sergeant, and Mariela Buonomo Zabaleta.

Contents

Contents

Contents

List of figures, tables, and boxes

Figures

Tables

Boxes

List of abbreviations

ACER	Australian Council for Educational Research
AFD	Agence Française pour le Développement
AU	assessment unit
BREDA	Bureau Régional pour l'Éducation en Afrique
BSD	Bureau de Stratégie et de Développement
CCM	Coordinating Committee Meeting
CDMT	Cellule de cadre de dépenses à moyen terme
CNA	capacity needs assessment
CREDD	Curriculum Research Evaluation and Development Directorate
DEF	Direction de l'enseignement fondamental
DFID	Department for International Development (UK)
DNEA	Directorate of National Examinations and Assessment
DNEF	Direction Nationale de l'Éducation Fondamentale
DPRE	Direction de la Planification et de la Réforme de l'Éducation
ECD	early childhood development
ECE	early childhood education
ECZ	Examinations Council of Zambia
EFA	education for all
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
EMIS	Education Management and Information System
ESA	education sector analysis
ESH	Évaluations Standardisées et Harmonisées
ESP	education sector plan

Abbreviations

ESSP	education sector strategic plan
FCDO	The Foreign, Commonwealth & Development Office (UK)
GABECE	The Gambia Basic Education Certificate Examination
GES	Ghana Education Service
GPE	Global Partnership for Education
IEF	Inspection de l'Éducation et de la Formation
IGE	Inspection Générale de l'Éducation
IIEP	International Institute for Educational Planning
ILSA	international large-scale assessment
INEADE	Institut National d'Étude et d'Action pour le Développement de l'Éducation
INRAP	Institut National de Recherche et d'Action Pédagogique
INSET	In-Service Education and Training
ISSEG	Institut Supérieur de Sciences de l'Éducation de Guinée
LSA	large-scale assessments
M&E	monitoring and evaluation
MEN	Ministère de l'Éducation nationale
MENA	Ministère de l'Éducation nationale et de l'Alphabétisation
MoBSE	Ministry of Basic and Secondary Education
MoE	ministry of education
MoGE	Ministry of General Education
NaCCA	National Council for Curriculum and Assessment
NAS	National Assessment Survey
NAT	National Assessment Test
NEA	National Education Assessment

Abbreviations

NEAU	National Education Assessment Unit
NGO	non-governmental organization
NIED	National Institute for Educational Development
NLAF	National Learning Assessment Framework
NSAT	National Standardized Achievement Test
OECD	Organization for Economic Cooperation and Development
PAD	planning and development
PADES	Programme d'Appui au Développement de l'Éducation au Sénégal
PAQEEB	Projet d'Amélioration de la Qualité et de l'Équité de l'Éducation de Base
PAQUET-EF	Programme d'Amélioration de la Qualité, de l'Équité et de la Transparence du secteur de l'Éducation et de la Formation
PASEC	Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN (Conférence des ministres de l'Éducation des États et gouvernements de la Francophonie)
PISA-D	Programme for International Student Assessment – for Development
PPARBD	Planning Policy Analysis Research and Budgeting Directorate
PQA	Programmes and Quality Assurance
ProDEG	Programme Décennal de l'Éducation de la Guinée
REXO	regional examination official
RTI International	Research Triangle Institute International
SABER	Systems Approach for Better Education Results
SDG	Sustainable Development Goal

Abbreviations

SEACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SIP	school improvement plan
SNERS	Système National d'Évaluation des Rendements Scolaires
SNESCO	Service National des Examens Scolaires et Concours
SNESE	Service National de l'Évaluation du Système Éducatif
SQAD	Standards and Quality Assurance Directorate
SSME	snapshot of school management effectiveness
STEM	science, technology, engineering, and mathematics
TALENT	Teaching and Learning: Educators' Network for Transformation
TIMSS	Trends in International Mathematics and Science Study
UIS	UNESCO Institute of Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations International Children's Emergency Fund
USAID	United States Agency for International Development
ZEEP	Zambian Education Enhancement Project

Executive summary

Background

For many students around the world, schooling is producing little or no knowledge, a state of affairs often referred to as the global learning crisis. The situation is particularly acute in sub-Saharan Africa, where nine out of ten children are unable to read and understand a simple text by age 10. Countries have made considerable progress in the development of their learning assessment systems, and large-scale assessments (LSAs) more specifically. Nevertheless, all too often, the use of learning data is still embedded in a linear understanding of the use of evidence, trusting that, as long as good quality data are available, there is a high chance that governments will actively consider them and make informed decisions accordingly. However, the implementation of assessments alone is not sufficient to trigger the use of learning data and, in turn, improvements in education quality. It is therefore essential to better understand the underestimated complexities underlying the use of assessment data, and to examine factors that facilitate or impede their use for planning. The study focused on national, regional, and international large-scale learning assessments, which are often introduced to provide education systems with a snapshot of learning achievements and inform policy. The project paid particular attention to actors' interactions, values, and perceptions when explaining the use of learning data, as these factors have received less attention in the available literature.

Executive summary

Research questions and methodology

The study was implemented in six sub-Saharan African countries (The Gambia, Ghana, Guinea, Namibia, Senegal, and Zambia). This report presents our findings and addresses the following research questions:

- What is the intended use of learning assessment data as defined in regulatory frameworks, and are different stakeholders aware of it?
- How do countries (i.e. ministry of education officials at different administrative levels) use learning assessment data in different phases of the planning cycle?
- What factors lead to the observed use of learning data?

The research in the project has been qualitative, which enabled us to better examine how different processes lead to the observed use of learning assessment data. The UNESCO International Institute for Educational Planning (IIEP) team, together with national researchers, conducted over 200 semi-structured interviews and multiple focus group discussions with ministry of education officials directly involved in the production and use of learning data at central and decentralized levels (e.g. assessments teams, planning units), as well as international partners and civil society representatives supporting assessments. We also collectively conducted direct observations of joint events at which national and international partners gather for policy dialogue or monitoring and evaluation (M&E) activities, as well as secondary data analysis of available documentation.

Main findings

The use of assessment data in the education planning cycle remains limited. Assessment data only sporadically inform planning processes and a narrow range of education policies:

- In education sector analyses (ESAs), the presentation of assessment data often remains descriptive and focuses on illustrating generally low learning levels. The analysis of the causes of poor learning outcomes is often missing or limited.
- It is difficult to establish a clear link between learning assessment results and the education sector plan (ESP) preparation. Learning assessment data often guide general orientations but, with a few exceptions, it is difficult to demonstrate their relationship to specific ESP policy measures.
- In the implementation phase, learning assessment data are mainly used by decentralized-level officials in the monitoring of school performance, development of school strategies, and organization of teachers' professional development activities.
- Learning assessment data most often inform the M&E phase of the ESP. They are used to set targets and report against them in ESP M&E frameworks and other M&E documents.

Some learning assessments end up serving certain purposes better than others. Early grade reading assessment/ early grade mathematics assessment (EGRA/EGMA) results had more impact on specific policy formulation,

especially in early grade literacy and numeracy, compared to other assessments. National LSAs more often served the purpose of diagnosis, whereas regional and international assessments compared the performance of different countries. National examination data are widely used by decentralized-level officials; they complement other assessments in the analysis of education quality, although they are less suitable for diagnostic purposes.

Learning assessment systems suffer from a lack of, or ineffective, regulatory frameworks to coordinate assessments and clarify their objectives. There is often no strategic national vision for learning measurement; instead, we observe a patchwork of assessments whose objectives in some cases overlap. The fact that different LSAs (i.e. national, regional, international) have been introduced gradually, and often as part of external projects, might have delayed the development of national regulatory frameworks.

Our findings also confirm other important factors (e.g. capacities, dissemination, funding issues) which have been explored in the available literature. They indicate that dissemination modalities are rarely defined strategically. In addition, high vulnerability to budgetary constraints often adversely affects the timely dissemination of learning data (especially at decentralized level). Assessment reports are often the only means of sharing LSA findings, whether national, regional, or international, in addition to dissemination meetings; however, national assessment reports frequently do not provide an in-depth analysis of key factors affecting learning, or targeted policy inputs, leading to challenges for decision-makers in extracting conclusive lessons.

Executive summary

Although international partners have played a key role in the development of learning assessments through financial and technical support, the effects of their support vary. In certain settings, their support has resulted in mutually beneficial relationships with national counterparts and strengthened assessment systems. In others, their influence has led to a power imbalance between different actors. Dependence on external funding sometimes leads to a lack of sustainability of assessment programmes, and limited national autonomy in assessment-related decisions. In addition, conducting large-scale learning assessments and ensuring effective use of their data are resource-intensive tasks requiring strong financial, human, and technical capacity that is often missing in the countries studied. This sometimes leads to the outsourcing of one or several tasks in the management of assessment data, leading to poor national ownership of the data produced. In some cases, learning assessments do not fully respond to national stakeholders' information needs, whereas in others, national actors perceive assessment data to be primarily intended to inform international partners' activities.

At central level, collaboration and communication among actors in relation to assessment results often remain sporadic. A dichotomy between the management of the political and technical dimensions of the assessment results is an important obstacle to the use of assessment data. Actors can sometimes overemphasize the technical roles of assessment teams, disconnecting them from decision-making processes.

Collaboration is often problematic between central and decentralized levels, as the management of assessment data frequently follows top-down dynamics. A feedback loop that would provide decentralized-level officials with information on assessment findings and expected follow-up action is often absent. Consequently, we observed a lack of awareness and low use of assessment data at this level.

Recommendations

Our findings lead us to several reflections and guiding principles to consider when looking for ways to improve the use of learning data:

Ministry of education leadership

- Clearly define the goals of your assessments and plan for the use of data when developing assessments.
- Gear the system towards national leadership and ownership.
- Commit to the development of national capacities in the production and use of learning data at different administrative levels.
- Reflect on the best institutional position of your assessment team.

National assessment teams

- Based on a national vision for assessments defined collectively, choose assessment design options that respond to identified actors' needs.

- Facilitate the understanding of learning data for different actors by adapting dissemination products and their content.
- Reporting is not the end of an assessment cycle: link dissemination activities with constructive feedback loops.
- Involve other actors in the management of assessments and look for synergies with other information sources, especially examinations.
- Adjust assessment cycles to regular planning and budgeting activities.

International partners

- Invest in capacities and transfer expertise, discourage outsourcing.
- Ensure that your support of learning assessments is in line with priorities defined in ESPs and other national strategic documents.
- Rely as much as possible on existing national assessments rather than creating new ones.

Introduction

Solving the global learning crisis: promises of learning assessments

In low and middle-income countries, 53 per cent of 10-year-old children cannot read and understand a simple story (World Bank, 2019a). For many students around the world, schooling is producing little or no knowledge, in addition to many children still being out of school, a state of affairs often referred to as a global learning crisis. This is particularly acute in sub-Saharan Africa. According to World Bank data (2019a), nine in ten children – that is, more than 100 million students – are learning poor[1] in the region, which is the highest regional percentage in the world.

The information provided by large-scale assessments (LSAs) is critical to understanding and addressing the global learning crisis in numerous ways. Although they are not the only metrics for evaluating the quality of education, student learning outcomes are key to estimating the extent of the learning crisis and the overall performance of education systems, which is why their measurement is high on the global education agenda. Assessment data are critical to monitoring progress towards the achievement of Sustainable Development Goal (SDG) 4.[2] They are also important to identifying education sector challenges and supporting the design and implementation of appropriate policies. When accurate, comparable, and up-to-date learning assessment

1 'Learning poverty means being unable to read and understand a simple text by age 10. This indicator brings together schooling and learning indicators: it begins with the share of children who haven't achieved minimum reading proficiency (as measured in schools) and is adjusted by the proportion of children who are out of school (and are assumed not able to read proficiently)' (World Bank, 2019a: 6).

2 SDG 4: 'Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.'

data is available, it can be used in educational planning and management to increase systems' efficiency, foster political engagement, and ensure equity (World Bank, 2018; UIS, 2017). These data also constitute valuable evidence that can contribute to the advancement of education research. Assessments can serve several of these purposes.

Complexities underlying the use of learning data

However, learning assessment data cannot fulfil their promise to improve education quality unless they are effectively used. Although more learning data has become available in low-income countries,[3] this seemingly has not led to their effective use in educational planning. Few research studies investigate the latter, but Saito and van Cappelle (2010) and Levine (2013), found that assessment data were not used to their full potential to inform educational planning and policy-making processes. A Global Partnership for Education (GPE) study reveals that only 24 per cent of the reviewed ESPs specified that some kind of evidence was used to identify the underlying causes of poor learning outcomes (GPE, 2017). These findings confirm that the adoption of learning assessments alone is not sufficient to trigger the use of learning data and, in turn, improvements in education quality. Hence, our interest is in investigating factors that support and impede the use of learning assessment data.

Much of the debate around learning assessments and the use of their data is still embedded in a linear understanding of

3 Middle- and low-income countries are now leading the growth of LSAs through their increasing participation in national and regional assessments (Tobin, Nugroho, and Lietz, 2016). In sub-Saharan Africa, 40 countries out of 46 have reported participation in at least one large-scale assessment (UIS, 2019). The same data show that 28 national learning assessments are conducted in the region.

the use of evidence in policy-making (i.e. assuming that the availability of knowledge will drive its use), which simplifies the realities of the issue. Discussions often focus on assessment validity and reliability as well as data comparability – in other words, technical elements of assessments. Although these elements are important, this leads to an over-emphasis on data quality when examining the use of data, trusting that, as long as good quality data are available, there are high chances that governments will analyse the results and make informed decisions accordingly.

In reality, many factors disrupt this linear pathway of the use of learning data. Firstly, the implementation of assessments and the use of their data is more challenging in contexts like sub-Saharan Africa, where countries are exposed to the same pressure to align with global metrics (e.g. reporting on SDG 4 indicators) as more developed countries, despite limited financial and human resources and capacities. Secondly, beyond the use of learning data to provide evidence for planning, participation in LSAs is driven by a range of factors that may pertain to the process itself (i.e. enhancing international relations, technical capacity development, receiving funding and aid) (Addey and Sellar, 2019). Thirdly, policy-making and planning are not linear; nor are they purely technical processes in which problems are identified, solutions considered, and the most appropriate strategies selected given the available evidence. Finally, actors' interactions shape not only the production but also the use of learning data. It is therefore essential to go beyond the linear understanding of the use of learning data and unveil the dynamic causes of their low use in sub-Saharan Africa if they are to contribute to efforts to improve education quality.

Introduction

About the project

Starting point and research questions

Development of learning assessment systems and improving use of their data is high on the agenda of many governments in sub-Saharan Africa. Whereas countries have already made considerable progress in the development of their learning assessment systems, the linkage between their data and educational planning is unclear. Despite the increasing attention that this topic is receiving, a deeper understanding of what is driving the use of learning data for educational planning is still missing.

The literature has identified several challenges that explain the low use of learning data. Reliability and relevance of information, financial and technical capacities, as well as coordination and dissemination channels, proved to be factors affecting the use of learning data (Raudonytė, 2019). Although important, they might be only one part of the explanation. Learning assessments are not just a technical tool but also a political phenomenon that reflects power relations between different stakeholders with diverging demands (Benveniste, 2002). Looking at this issue from different actors' perspectives seemed promising. How do actors perceive the intended use and usefulness of learning data? How do they interact, and how does this affect the use of learning data? Finally, how do different factors interact with and shape the use of learning data?

The need to address this gap prompted the UNESCO International Institute for Educational Planning (IIEP), with its longstanding experience and expertise in educational

Introduction

planning, to launch a research programme to examine how and why countries in sub-Saharan Africa use learning assessment data in the education planning cycle, paying particular attention to analytical approaches from political economy and institutional analysis.[4] The political economy approach was used to understand how diverging stakeholders' interests and incentives, which are themselves a result of the country's broader institutional, social, and political context, influence the use of learning data. Institutional analysis perspectives focused on the examination of how countries' educational administrations functioned in the areas of strategic planning, policy design and implementation, management of information systems, and human and financial resources when it comes to the use of learning data. These factors were expected to have an influence on the use of learning data and were therefore selected for the analysis.

In conducting this study, IIEP intends to provide ministries of education (MoEs) as well as development partners with a comprehensive understanding of how to sustainably improve the use of learning data. It also seeks to raise awareness about the complex dynamics that are involved in the use of learning data and that have often been overlooked in the literature, especially with regards to educational planning. Indeed, the use of learning assessment data in the planning cycle has not received a lot of attention so far. Moreover, few studies explore the use of data generated by LSAs in sub-Saharan Africa.

Yet the education planning cycle has become a prominent sequence of processes in many developing countries. GPE has

4 It is important to underline that both are closely linked, as political economy analytical frameworks often include some elements from institutional analysis.

Introduction

made the development of credible and robust sector plans a mandatory condition for countries' aid eligibility. Moreover, grant requirements and processes (e.g. ESP development and appraisal, grant implementation, and monitoring) are run alongside the planning cycle (GPE, 2020). This progressive standardization of planning processes, especially in sub-Saharan Africa,[5] makes it particularly relevant to study the use of learning assessment data in the education planning cycle in the region.

The following research questions thus underpin the project:

- What is the intended use of learning assessment data as defined in regulatory frameworks, and are different stakeholders aware of it?
- How do countries (i.e. MoE officials at different administrative levels) use learning assessment data in different phases of the planning cycle?
- What factors lead to the observed use of learning data?

However, study findings presented in this book go beyond direct responses to initial research questions. The research project allowed us to collect rich information on the use of learning data and factors that influence it that exceed the initial scope of the study. Firstly, results go beyond the scope of the planning cycle and include considerations for broader policy dialogue and policy-making. Secondly, it involves analysis of a wide spectrum of factors that directly or indirectly affect the use of data. Some of those factors go beyond the initial pre-identified areas of analysis and emerged from the data exploration (including literature sources), namely:

5 In 2019, almost three-fourths of the total GPE implementation grant funding was disbursed to sub-Saharan Africa (GPE, 2019a).

- What are the potential uses of learning assessment data in the planning cycle phases? (*Chapter 1, Section 1.2*)
- What types of policies do learning assessment data inform? (*1.3.1*)
- What is the comparative influence of different learning assessments? (*1.3.2*)
- What are the advantages and disadvantages of different institutional settings of assessment teams, and what did we find in the project countries? (*2.4*)
- How can an assessment design affect the use of learning assessment data? (*2.3*)
- What are the preconditions for more effective assessment policies? (*2.2.3*)
- How are assessment data associated with results-based schemes and what implications does this have? (*5.4*)
- What role do learning assessment data play in capturing external funding? (*5.5*)

Research process

The study began in 2018 with an in-depth review of the existing literature to identify research gaps. Its main findings were published in the IIEP Working Paper *Use of Learning Assessment Data in Education Policy-Making* (Raudonytė, 2019). It reviewed how learning data inform policies and the most common barriers and risks to the use of learning data. To address identified research gaps and explore the use of learning data in the planning cycle, focusing on actors' dynamics, it was followed up with field research. Research tools were piloted at the end of 2018 in Guinea and The Gambia. They were then slightly modified and field research continued throughout 2019 and in early 2020 in the remaining four countries of the study (Ghana, Senegal, Zambia,

Introduction

Namibia). The IIEP project team and national researchers carried out data collection collectively. The IIEP team was mostly involved in data collection at central ministry level, often together with national researchers, who then continued data collection at the first decentralized level (i.e. region/province). However, these modalities varied slightly from country to country. The IIEP team, together with national researchers, conducted semi-structured interviews and focus group discussions with MoE officials and international partners working on the production and use of learning data; we collectively conducted direct observations of joint events that brought national and international partners together for policy dialogue or monitoring and evaluation (M&E) activities. A total number of 227 interviews, 22 focus group discussions, and 10 observations were conducted in this study (please refer to *Annex 2* for more information on data collection modalities in each country, as well as a more detailed presentation of the study methodology). The IIEP team also conducted secondary data analysis of documents available at country level.[6] The combined use of primary and secondary information and the multitude of actors that were consulted increased the overall consistency of the analysis.

The IIEP research team then proceeded with an in-depth examination of collected data and preparation of country-level analyses. Deductive coding was used to analyse qualitative data as per the key factors the study wanted to examine, as identified in its analytical framework (see below). However, some additional themes have also emerged when analysing

6 These are: education sector plans, education sector analyses, budget reports or evaluations, M&E documents, national and international studies, education laws, relevant policy documents/regulations, external programming documents, other relevant documents (e.g. PowerPoint presentations for specific events).

the data. The research team then released the main findings for each country in policy briefs and information sheets.[7]

Depending on the setting, national researchers were national university employees, MoE officials, or independent consultants. Their knowledge of the national context was crucial to contextualizing the study in countries, identifying relevant participants and facilitating exchanges with them. To ensure researchers understood the project and its methodology, they received 'IIEP Research Guidelines' with a detailed research protocol, research tools, and specific guidelines on how to use them. The IIEP project team also facilitated face-to-face or online training discussing common challenges encountered in the pilot study implementation.

The research in the project is qualitative, which allowed us to better examine how different processes lead to the observed use of learning assessment data. Its multi-case study design allowed for the analysis of country specificities while at the same time highlighting common challenges and achievements, and the complexity that lies in the interactions of different factors influencing the use of learning assessment data.

Analytical framework

The analytical framework developed at the inception of the study guided the data collection and analysis. It entailed four main sections with specific analytical points to study, mirroring the main research questions:

7 http://www.iiep.unesco.org/en/our-expertise/use-learning-assessment-data.

Introduction

- Institutional architecture for managing learning assessment data, exploring regulations on learning assessment systems, their goals, the division of roles and responsibilities among actors involved in the management of learning data, as well as strengths and weaknesses of the formal setting.
- Actors' knowledge of existing regulatory documents on the use of learning data and their understanding of their official intended use.
- Their knowledge of actual practices which is linked to the use of learning assessment data in the planning cycle.
- Their insights on factors that determine the way these data are used, focusing on political economy and institutional analysis elements.

The last section included a number of factors to analyse that were grouped into the following subsections:

- Historical legacies – countries' past experience with learning assessments and patterns of historical use that may have effects on the current use of data.
- Ideologies and values – actors' narratives and perceptions regarding the use of learning data.
- Enforcement and functioning of formal institutions – degree of coordination and cooperation between structures in charge of the management of learning data, material and human resources, as well as autonomy to implement assessments and national officials' capacities.
- Informal institutions – informal social and cultural norms and practices that shape power relations and planning processes, and which might influence the use of learning data.

— Actors and power relations – the balance of power between different actors and 'pockets' of resistance and support for the use of learning data.
— Other potential barriers and enabling conditions (e.g. data dissemination channels, timely dissemination, quality of assessments and their reports, coherence with other policies and data sources).

These subsections were identified and selected based on an in-depth review of existing frameworks for analysing political economy in the education sector, as well as studies that documented the use of learning data in other contexts (i.e. other regions or policy areas). This comparative analysis provides an in-depth analysis of factors that proved important in the study and on which we managed to collect a rich set of information.

Country selection

The study covered sub-Saharan Africa because it is UNESCO's regional priority, and because their national assessment systems have received less attention in the literature. The study was implemented in six countries: The Gambia, Ghana, Guinea, Namibia, Senegal, and Zambia (see *Figure 1*).

The following criteria were applied when selecting countries:

— They had functioning student assessment systems in place, as this is the first prerequisite for the use of learning data. The study included countries that have improved the use of learning data over time, as well as those facing challenges when using that data. This allowed a range of

Introduction

different country situations to be covered, and to identify success stories and emerging challenges alike.

— The selection included a number of countries from (a) SEACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality) assessment; (b) PASEC (Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN) assessment; (c) countries not involved in any regional or international assessments but with functioning national assessments.

— IIEP also favoured countries in its close network, thus facilitating initial contacts with researchers and officials in MoEs.

Figure 1. Countries selected for the study

Source: Developed by authors.
Note: The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of UNESCO or IIEP concerning the legal status of any country, territory, city, or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. Final boundary between the Republic of Sudan and the Republic of South Sudan has not yet been determined. Final status of the Abyei area is not yet determined.

Scope

The study considered national, regional, and international large-scale assessments (LSAs), as well as early grade reading assessment (EGRA)/early grade mathematics assessment (EGMA) and citizen-led assessments. These are low-stakes assessments for students and teachers as they do not influence students' education trajectory, but rather provide a snapshot of learning achievement for a group of learners in a given year and in a limited number of subjects. Among other elements that the book explores, we try to understand whether the uptake of learning assessment data differs by type of assessment. *Table 1* presents assessments conducted in project countries that were included in the analysis.

Our findings focus on basic education and do not cover adult literacy, technical and vocational education and training, higher, or non-formal education. Moreover, the study did not examine the use of learning data by teachers.

When speaking about the *use* of learning data, the study refers to the process by which policy-makers and planners actively consider, analyse, and engage with these data when conducting planning activities (building on Porter's [2010] definition of the use of research). It includes, but is not limited to, decision-making in line with assessment findings.

The study examines the use of learning data in the education planning cycle and educational planning, which 'is the application of rational, systematic analysis to the process of educational development with the aim of making education more effective and efficient in responding to the needs and

goals of its students and society' (Coombs, 1970). *Chapter 1* provides further details of how the study conceptualized the planning cycle. Nevertheless, in many cases, a broader view that expanded to education policy was adopted, especially for discussions with actors less engaged in planning activities.

Acknowledging that assessment data are part of an information ecosystem and cannot be interpreted alone, since this would only provide a limited analysis of the education system issues, an analysis that may be misleading, the study also analysed how learning assessment data were articulated with other assessment results (e.g. examinations, classroom assessments) and other available evidence derived from the Education Management and Information System (EMIS) and national/international studies.

Limitations

It was difficult to examine dynamics among actors as well as patterns of data use during observations. By design, the exercise implies a certain degree of subjectivity on the researchers' side. There was also a risk of a 'reactive effect' when research subjects modified their behaviour knowing that they were being observed, or adapted their discourse depending on what they thought was expected from them (Bryman, 2004).

Table 1. Assessments conducted in project countries

	The Gambia	Ghana	Guinea	Namibia	Senegal	Zambia
National assessments	National Assessment Test (NAT)	National Education Assessment (NEA)	Evaluation Nationale des Acquis	National Standardized Achievement Test (NSAT)	Système National d'Evaluation des Rendements Scolaires (SNERS)	National Assessment Survey (NAS)
Regional assessments			PASEC	Southern and Eastern Africa Consortium for Monitoring Educational Quality (SEACMEQ)	PASEC	SEACMEQ
Hybrid assessments	EGRA[8]	EGRA	EGRA	EGRA	EGRA	EGRA
	EGMA	EGMA				EGMA
International assessments	Trends in International Mathematics and Science Study (TIMSS)				Programme for International Student Assessment – for Development (PISA-D)	PISA-D
Citizen-led assessments					Jàngandoo	

Source: Compiled by authors.

8 EGRA and EGMA have been classified as either international (Raudonytė, 2019) or hybrid assessments (Wagner, 2011).

Although the study aimed to observe events that gathered national and international actors together for policy dialogue or M&E activities, due to time constraints we could not participate in certain events, thus influencing the selection of events for observation. In addition, conflicting schedules made the organization of focus group discussions difficult, as many participants could not attend at the same time.

Ultimately, a potential bias could have arisen from the IIEP's involvement in the preparation of Guinea's ESP because the study was implemented simultaneously. Preliminary study findings might have influenced the development of the plan. In addition, the United States Agency for International Development (USAID) runs country projects and finances some learning assessments in Ghana and Senegal, where its consultant participated in data collection at central level.

Objectives and structure of the book

As an increasing number of low-and middle-income countries conduct LSAs, this book aims to provide national planners, policy-makers, and development partners with new evidence on how to support more effective use of learning assessment data in educational planning. It reports on the main findings of the study and aims to provide a more comprehensive and analytical understanding of the issue.

The book is structured as follows: *Chapter 1* examines how learning assessment data are used to inform the planning cycle and selected education policies (research question 2). *Chapter 2* explores institutional settings for the use of learning data (research question 1). The rest of the book focuses on factors influencing the use of assessment data

(research question 3). *Chapter 3* sheds light on the analysis of selected barriers to the use of learning data in sub-Saharan Africa, focusing on their interactions. *Chapter 4* explores the interplay between national actors at different administrative levels and how this affects the use of learning assessment data. Finally, *Chapter 5* focuses on the role of international partners in the development of assessment systems and the influence this has on the use of learning data. *Annex 1* provides definitions of the main terms used in the document, whereas *Annex 2* details the methodology used in the study.

Chapter 1. Overview of the use of learning assessment data in planning cycle phases

The IIEP/GPE *Guidelines for Education Sector Plan Preparation* (IIEP-UNESCO and GPE, 2015) recognize that one of the essential features of a credible sector plan is the use of data and assessment information in its development. This chapter defines the planning cycle and its phases. It then reports the observed uses of learning assessment data in different phases of the planning cycle, and examines policies informed by these data.

1.1 How can learning assessment data inform the planning cycle?

1.1.1 Definition of a planning cycle

Our conceptualization of the planning cycle broadly mirrors policy-making phases.[9] Educational planning follows cyclical processes. It encompasses the following key stages that are formalized into strategic programming documents (see *Figure 2*):

1. Education sector analysis (ESA): analysing the current situation in the sector.

2. Education sector plan (ESP) preparation: defining precise objectives and programmes.

9 A simplified conceptualization of a policy cycle consists of the following phases: (1) Agenda setting: raising awareness and prioritizing an issue. (2) Policy formulation: selecting and constructing options and strategies. (3) Policy implementation: determining the practical activities to be implemented. (4) Monitoring and evaluation: monitoring and assessing the process and impact of an intervention (Sutcliffe and Court, 2005: 5–6).

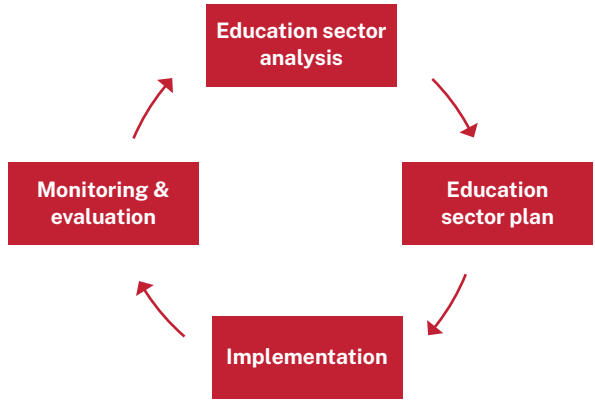
1. Overview of the use of learning assessment data in planning cycle phases

3. Implementation: annual planning through operational plans and budget updates.

4. Monitoring and evaluation (M&E): measuring progress for corrective action.

These phases do not always follow a perfectly cyclical pattern, as the processes are iterative and sometimes overlap. Moreover, for conceptualization purposes, a simplified version of the planning cycle was purposely chosen. A more comprehensive version can be found in the IIEP *Guidelines for Education Sector Plan Preparation* (IIEP-UNESCO and GPE, 2015: 13).

Figure 2. Education planning cycle



Source: Adapted by authors from IIEP-UNESCO and GPE, 2015.

Learning assessment data have the potential to inform each step of the planning process. The next section explores potential specific channels for doing so.

1.1.2 The potential use of learning assessment data in the planning cycle phases

Educational planners and other actors in the education system systematically need to know whether students are learning. Are there learning inequities within the education system? What are the main obstacles to learning? What priorities and strategies can effectively overcome these obstacles? Learning assessment data have a lot of potential to provide responses to these questions throughout the education planning cycle.

Education sector analysis

An ESA is the first step in sector planning; learning assessment data can serve multiple purposes in the in-depth and holistic diagnosis of the education system:

- Learning data provide information on student learning outcomes and quality of education.
- By identifying gaps in learning, assessment data can shed light on equity issues (regional, gender, linguistic, socio-economic, etc.), raise awareness around them, and provoke debate, bringing them to the discussion table.
- The information from background questionnaires administered together with assessments, and even administrative data in some cases, can be associated with student results to provide insights on teacher, school, and student-related factors that might influence learning.
- Learning assessment data can be used to develop a cost-effectiveness analysis that considers the effectiveness of different interventions to improve learning in relation to the resources used.

1. Overview of the use of learning assessment data in planning cycle phases

Education sector plan

An ESP 'provides a long-term vision for the education system in the country and outlines a coherent set of practicable strategies to reach its objectives and overcome difficulties' (IIEP-UNESCO and GPE, 2015: 9). Learning assessment data can provide a rich set of information on the key determinants of learning, as well as subjects, geographical areas, or population groups that deserve policy-makers' attention. They can inform the overall vision presented in ESPs, as well as specific strategy formulation.

Implementation

The next step in the planning process consists of defining what practical measures are needed to translate the ESP vision into the implementation of activities. The *Guidelines for Education Sector Plan Preparation* explain that the plan 'outlines the detailed activities for a specific period of the plan [typically a medium-term period], with information on timing, roles, responsibilities, and costs' (IIEP-UNESCO and GPE, 2015: 23). Learning data might be particularly helpful for actors in charge of implementing the plan at subnational level:

- They can be useful in school-level management for setting and monitoring school-level goals and informing school improvement plans.
- Relevant authorities can consider assessment results in their decisions regarding teacher support, professional development, and postings.
- Learning assessment results can inform parents interested in schools' performance.

1. Overview of the use of learning assessment data in planning cycle phases

Monitoring and evaluation

The effective implementation of an ESP requires M&E mechanisms to ensure activities are carried out as intended, and targets are achieved (IIEP-UNESCO and GPE, 2015). An ESP M&E framework includes the expected results of the plan, indicators to measure their achievement, and targets set for the period of plan implementation, as well as the data sources and parties responsible for their production (IIEP-UNESCO, 2019b). Learning assessment data can inform the achievement of the ESP learning targets and evaluate the effectiveness and efficiency of specific ESP strategies. Beyond the ESP, these data can also be used to report on the education sector's overall performance to national stakeholders and international partners.

Risks of misusing learning data

The use of learning data entails numerous risks – documented in the literature – that also need to be considered in this analysis. Firstly, methodological issues can adversely affect data accuracy and result in misleading interpretation. Learning assessments are highly technical endeavours requiring strong technical capacities that are not always available in countries. Secondly, when assessment data are used in isolation, with little or no consideration for other evidence (e.g. qualitative studies, administrative data on teachers), there are risks that the complete picture has not been properly examined as a whole. This might misinform the interpretation of assessment results. In addition, a number of contributions have shown that assessment data can be used to legitimize predefined agendas, rather than to construct policies informed by evidence. Finally, as assessments

are conducted on a small number of subject areas, over-reliance on assessment data can narrow our understanding of education to a limited number of measurable indicators and over-simplify a complex reality. Finally, all too often, correlations between assessment results and factors affecting learning are interpreted as causal relations, which also misinforms data interpretation (Raudonytė, 2019). Although this study does not go further in exploring these risks in depth, it is important to consider the risks involved in the use of learning data when analysing their use, and to interpret our study findings accordingly.

These are multiple ways through which learning assessment data can feed into each of the planning cycle phases, and there are numerous risks when using them. However, how do educational planners use these data in practice? The analysis will now turn to the *observed uses* of learning assessment data in the planning cycle in the countries studied.

1.2 The use of learning assessment data in the planning cycle: a loss of information throughout the process

Although in sub-Saharan Africa countries there have been a growing number of references to learning assessments in ESPs since the early 2000s (Furiv, unpublished manuscript), our study results demonstrate that the use of assessment data in the education planning cycle is somewhat limited, as they inform planning processes sporadically. They also attest to a loss of information along the cycle and a heterogeneous use of data in different planning phases.

1. Overview of the use of learning assessment data in planning cycle phases

1.2.1 The analysis of learning data in ESAs provides limited information to inform ESP strategies in a meaningful way

Discussion on learners' performance often remains very general

An ESA is often a reference document in ESP preparation discussions. If an ESA does not capture the main messages and the richness of information provided in learning assessment reports, this can contribute to poorer mobilization of learning data in ESPs. Our study therefore examined how learning data appear in ESAs (or equivalent documents with similar features and purposes) in the countries studied.

> Learning assessment data have the potential to inform each step of the planning process.

Although all ESAs present learning assessment data and examination results and disaggregate information to illustrate overall levels of learning, this presentation often remains descriptive. Nevertheless, it provides relevant elements that are an important first step in the analysis. Elements often covered in ESAs include the following:

- Assessment results are frequently displayed per proficiency levels, although these are not always clearly defined or fine-grained enough to allow for a more detailed visualization of scores' distribution.
- Most ESAs present a comparison of students' results over time, generally between the last two or three assessment cycles.
- Assessment results are often disaggregated by regional, urban/rural, socio-economic status, or gender variables.

1. Overview of the use of learning assessment data in planning cycle phases

> Disaggregating data allows learning gaps to be highlighted across subgroups of the population, which could contribute to guiding equity policies.

Although these elements provide important first insights into the learning situation, the discussion then often remains at a very general level, focusing on illustrating generally low learning levels (with exceptions as outlined below). This does not always lead to a clear identification of challenges and progress achieved. An extract from the Namibia ESA is illustrative:

> Average scores in Standardized Achievement tests at Grade 5 level remained below 50% from 2009 to 2014. However, in 2015, there was an encouraging step-up in Grade 5 scores in both English and Mathematics. If sustained, this improvement could herald the achievement of better-quality results throughout the system in the years to come. There has been a gradual improvement in Grade 7 SATS Mathematics scores, but grade 7 English scores declined in 2015. (Namibia, 2017: 8)

Likewise, the section on learning achievements in the Senegalese ESA only provides global scores obtained in PASEC 2014 and the national learning assessment (i.e. Système National d'Évaluation de Rendement Scolaire, SNERS) with no further description or analysis.

1. Overview of the use of learning assessment data in planning cycle phases

Analysis of factors influencing learning in ESAs is a missing link

Learning assessments are usually administered along with background questionnaires that gather input information on in-school (e.g. teachers' educational background) and out-of-school (e.g. parents' literacy) factors, which can enrich the interpretation of students' scores. Statistical techniques such as econometric regressions are common tools for analysing the relationships between these inputs and outputs. Although caution must be applied when drawing inferences from regression results, they can be particularly helpful to pinpoint key factors likely to affect learning and requiring further attention or investigation.

Investigating the causes of poor learning outcomes is critical to informing refined policies, but our study demonstrated that this analysis in ESAs varies by country. It is either limited or omitted in four ESAs under our review (Namibia, Senegal, Ghana, Zambia) because they rarely perform advanced statistical tests to explore the roots of low learning achievements. However, the ESAs of Guinea and The Gambia provide examples of an in-depth analysis of factors affecting learning (see below). The following concerns emerged from the review of national ESAs:

- When factors influencing learning are discussed, they are often examined in isolation and dispersed across different ESA sections. This makes it more difficult to have a full picture of factors influencing learning and their relative importance.
- ESAs do not always contain annexes with technical details on the learning data analysis, i.e. detailed results of

statistical tests conducted to analyse learning assessment data. Therefore, it is sometimes difficult to interpret the results.

— In addition, across countries, some factors come up frequently in ESAs (e.g. teacher and student absenteeism, inadequate textbook provision, or poor-quality teacher training). Although these may be determinants of learning, it is difficult to ascertain whether they have also been identified through a country-specific in-depth analysis of assessment results. Indeed, references to sources are often missing, whether learning assessments or others.

Our review also found promising examples of an in-depth analysis of learning data that provides more insights for policy-makers. ESA authors in Guinea and The Gambia engaged with raw data and conducted an in-depth econometric analysis using statistical tools to assess key obstacles to learning. They also looked into the cost-effectiveness of some policies using learning data. *Box 1* provides more information on the analysis performed in the Gambian ESA.

> “ Investigating the causes of poor learning outcomes is critical to informing refined policies.

1.2.2 Inconclusive evidence on how learning data influence the development of ESP strategies

Our research sought to understand whether and how learning data inform education strategy formulation in the process of an ESP development.

1. Overview of the use of learning assessment data in planning cycle phases

BOX 1

Analysis of learning assessment data in The Gambia Education Country Status Report (2011)

General student scores in EGRA and a National Assessment Test (NAT) are reported; learning assessment data are disaggregated by region, gender, and socio-economic status. Multivariate analyses have been performed, using a logistical model, to better understand what factors have an impact on NAT scores. It is worth noting that the analysis lists all variables considered, including those that have not been found to be significant. The document also notes that other factors, which might not have been captured by the current model specification, can affect student results, and that additional in-depth analysis on learning outcomes and improved tools are needed to suggest clear policy orientations.

The ESA explores different policy implications of the NAT results. For instance, the authors analyse the relationship between the lower basic schools' endowment and their NAT Grade 5 performance. As this relationship is weak, it implies that better-resourced schools do not systematically perform better and, conversely, schools with poor results are not always underfinanced. In addition, available options for improving learning achievements, with their potential impact and cost, are evaluated.

The challenge of examining the influence of assessment data in the selection of specific education strategies

Examining the extent to which an ESP is informed by learning assessment data is an inherently challenging exercise, as it requires reconstructing a chain of processes that allows

1. Overview of the use of learning assessment data in planning cycle phases

evidence to be linked with chosen strategies, which is a difficult task for several reasons.

Firstly, even if some strategies are consistent with lessons from the analysis of assessment results, official documents rarely demonstrate an explicit reliance on learning assessment data or other evidence (citing sources used to inform a decision). Secondly, while many respondents reported that learning data were useful in the ESP preparation process, in all countries studied, many respondents were not able to indicate a specific programme that was informed by these data. On the contrary, it seemed that interviewees often expressed wishful thinking, referring to an ideal situation (in which evidence is used for policy-making) or the intended use of learning assessment data, rather than their actual use. Two comments from The Gambia reflect this well: 'Most of the strategies in the ESP are data-driven', and 'I want to believe that it has been used ….'.

Thirdly, policy-making processes are complex and often non-linear. They are influenced by different pieces of evidence and other factors, making ESAs and learning data just one element of information among others. As Addey and Sellar (2019: 6) put it, 'establishing a causal relationship between ILSA [international large-scale assessment] data and educational reforms is problematic, because policy processes are complex and rarely driven by a single causal factor'. Policy-makers' decisions are influenced by a number of other factors (e.g. experience, ideology, financial resources, current trends). In addition, our research led to the observation that not all policies are embedded in the regular planning cycle. Some policies influenced by assessment data were designed or implemented before/after the preparation of ESPs or as part of external projects in parallel to ESPs.

1. Overview of the use of learning assessment data in planning cycle phases

Learning assessment data guide general orientations in educational planning and policy-making

Although learning assessment data have brought attention to certain issues and influenced some specific education strategies, their use currently remains limited to broad considerations, as highlighted by a planning officer in Ghana: 'We do well at the aggregated level to inform macro policies. I am not sure we do well in going beyond aggregated data to do an analysis to inform targeted interventions. We don't do this very well.'

The following excerpt from the 'PISA-D Capacity-building Plan' for Zambia indicates the same issue about SEACMEQ's influence in the country:

> Although these existing research-oriented LSAs have a lengthy track record in Zambia, their direct impact on educational policy and educational discourse has been limited. For example, the SEACMEQ results have prompted additional scrutiny on the education sector, but there are no policies or initiatives that have been produced as a direct result of research findings of either assessment programmes. (OECD, 2016: 10)

Indeed, when asked about the influence of learning assessment data in policy decisions, research participants often reported that learning assessment data were used to inform general policy orientations and trigger discussions rather than inform specific programmes. In Namibia and The Gambia, the orientation of policies towards the expansion of early childhood development (ECD) has been encouraged by assessment results. A World Bank representative in The Gambia explained that EGRA results moved the policy focus to foundational skills in numeracy and literacy in the early grades.

1. Overview of the use of learning assessment data in planning cycle phases

Observations conducted in Guinea revealed that learning data were mentioned in general terms and sporadically throughout the preparation of an ESP, pointing to average scores to illustrate alarming proficiency levels, rather than used to advocate for specific policies. The example of Guinea is illustrative. During the two and a half days of discussions on the education quality strategies of the Ten-Year Education Programme (Programme Décennal de l'Éducation en Guinée [ProDEG]), stakeholders, including the assessment team, did not refer to specific findings of assessment reports beyond general references to the fact that learning outcomes were low (even though other studies were used).

1.2.3 Implementation of an ESP and the use of learning data at decentralized level

In the project countries, the use of assessment data in an ESP implementation is primarily driven by decentralized-level officials[10] and the extent to which they rely on these data. Although that use overall remains limited, our study documented some interesting examples. For instance, in The Gambia, learning data are heavily used for subnational level planning activities because this use is institutionalized.

Results of The Gambia NAT inform the participatory performance monitoring system, which aims to encourage the participation of parents and communities in education. The system consists of two components:

10 The study examined the use of learning data at the first level of decentralization (i.e. provinces/regions), except for Ghana and Senegal where some district level officials were also interviewed. All countries in the study are unitary states.

1. Overview of the use of learning assessment data in planning cycle phases

- School Participatory Performance Monitoring (SPMM), as an annual meeting that brings parents and communities together to discuss students' results;
- community record cards that capture students' performance in the NAT and The Gambia Basic Education Certificate Examination (GABECE) in each school, compared to the available resources and performance of other schools in the region/district/nationally (see *Figure 3*).

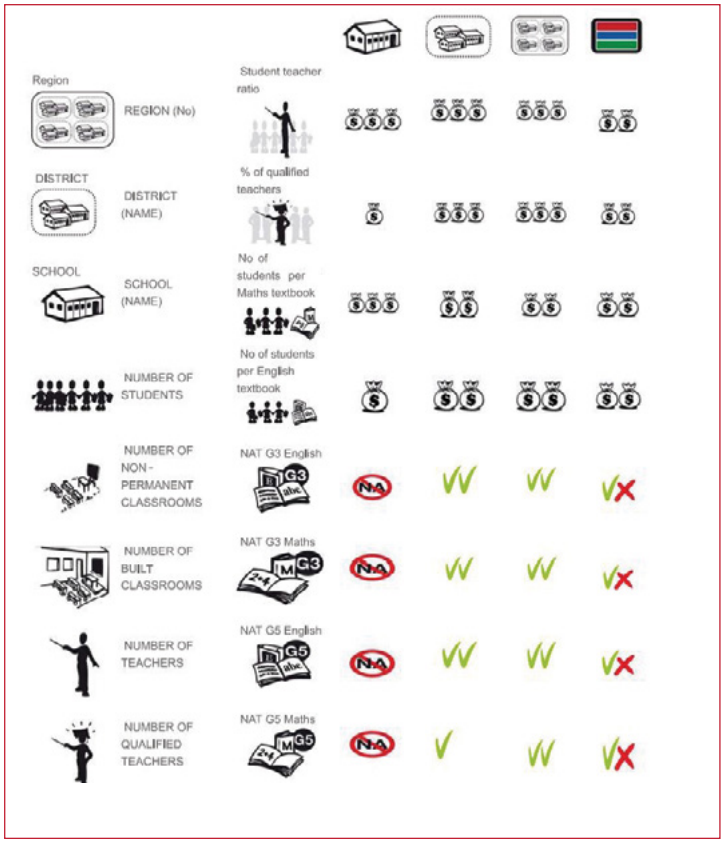
As part of a quality assurance mechanism, the NAT results are used to set targets and develop strategies in school improvement plans (SIPs) (i.e. school-level planning documents). School improvement grants are then allocated based on the quality of SIPs and NAT attendance.

In other countries such as Namibia and Zambia, regional or provincial-level officials use learning data to inform teachers' professional development activities. In Ghana, district officers refer to EGRA results, produced within the framework of the USAID-supported Learning Programme, to identify teacher weaknesses and proceed accordingly with professional development or replacement measures. Officers at decentralized level also rely on learning data to inform some of their planning activities and track the performance of schools, districts, or regions/provinces.

Many interviewees highlighted that the use of learning assessment data in planning at decentralized level strongly depends on the level of data disaggregation and their availability at that level. Since regional and international LSAs rarely meet these requirements, officials at subnational levels rarely use them. Instead, national assessments and EGRA/EGMA results are more frequently exploited.

1. Overview of the use of learning assessment data in planning cycle phases

Figure 3. Community card in The Gambia



Source: Ministry of Basic and Secondary Education (MoBSE) website (www.edugambia.gm/).

1. Overview of the use of learning assessment data in planning cycle phases

For example, in Senegal, assessments conducted by regional academies[11] are an important source of information for regional planners. In addition, in all project countries, examination data provide even more detailed and yearly information, which makes such data a more practical tool for some planning activities. Many of the planning activities at decentralized levels relate to day-to-day management of equipment, learning material provision, class size, etc., thus requiring timely student-or school-level data. For this reason, examination data are favoured over LSA data at decentralized level routine activities. Overall, examination results provide important information on student achievements and quality of education which can complement other information types.

Nevertheless, it is crucial to note that examination data are less suitable for diagnostic purposes. Firstly, there are no background questionnaires that could help interpret their score variation. Secondly, a detailed analysis of public examination results is not shared with students, parents, or teachers. The scores might be shared with schools, but they are not linked to the specific questions.

Our study also demonstrated that limited data analysis is taking place by the external public (e.g. national universities).

11 These are the Standardized and Harmonized Assessments (Evaluations Standardisées et Harmonisées [ESH]), Harmonized Progressions and Standardized Assessments (Progressions Harmonisées et Evaluations standardisées [PHARES]), and the Project for the Improvement of Basic Education Quality and Equity (Projet d'Amélioration de la Qualité et l'Equité de l'Education de Base [PAQEEB]) assessments.

1.2.4 M&E, the planning phase most informed by learning assessment data

To assess the use of learning assessment data in the M&E phase, our research examined indicators of student learning outcomes in ESPs' M&E frameworks and their use in related documents, such as joint sector reviews and mid-term/final evaluations of ESPs. Similarly to Tobin, Nugroho, and Lietz (2016), our findings suggest that M&E are more frequently informed by learning assessment data than by other planning cycle phases.

Indicators of student learning appear in all M&E frameworks that we examined. Assessment data inform them and help track progress in the implementation of ESPs. Indicators take various forms, as shown in *Table 2*, and they mostly rely on national assessments and EGRA/EGMA results. Interestingly, regional and international assessment data are rarely used to inform indicators in M&E frameworks; however, our study could not identify reasons for this.

Table 2. Types of indicators to report on student learning in M&E frameworks

Type of indicator	Countries
Percentage of students achieving proficiency level	Guinea, Ghana, Zambia
Percentage scored in the test	Namibia
Variation rate of student scores	Senegal
Variation rate of students passing the test	The Gambia

Source: Compiled by authors.

1. Overview of the use of learning assessment data in planning cycle phases

Nevertheless, several challenges persist. Firstly, some M&E frameworks do not explicitly refer to specific assessments or parties in charge of collecting and reporting these data. It was sometimes difficult to understand which assessment results would be used to inform indicators. For instance, the Guinea ESP (2015–2017) contains a single indicator on learning: 'Student achievement rates in reading in Grade 2' (Guinea, 2014a). It is not clear which assessment is used to inform this indicator, as both the national assessment and the EGRA assess the reading abilities of Grade 2 students. Secondly, while indicators' values should be updated from one ESP round to another, this has sometimes been difficult due to the discontinuity in assessment administration, impeding a regular update of indicators' values. In the Senegalese ESP (Programme d'Amélioration de la Qualité, de l'Équité et de la Transparence du secteur de l'Éducation et de la Formation (PAQUET-EF) 2013–2025), the variation in student scores has not been reported for any of the years considered in the plan. Thirdly, proficiency levels are not always defined. This was the case in the Guinea ESP (2008–2015), in which the indicator of learning was defined as the percentage of pupils achieving 'acceptable' proficiency in given subjects without further details on what the latter constitutes (Guinea, 2007).

The use of learning data to inform joint sector reviews and mid-term/final evaluations of ESPs remains limited. General scores or a descriptive analysis are briefly presented without further examination of the factors influencing learning. Assessment results are not used to evaluate the effectiveness of ESP strategies in these documents. The use of learning assessment results in annual or mid-term reviews to evaluate the effectiveness of the chosen strategies might indeed be

problematic, as it takes time for any learning strategy to show its impact through learning assessment results.

However, decentralized administrative levels are heavily involved in the M&E of education policies and frequently rely on learning data. In Senegal, school grants are conditional on learning results (i.e. as per contracts of performance between districts and schools in the framework of the PAQEEB project). In Namibia, annual review meetings are an opportunity for regional directorates to report on national progress, based on the average scores obtained in the NSAT. However, observations conducted at these meetings showed that only general scores were presented, without correlations linking student results and influencing variables. The discussion was instead more oriented towards administrative issues.

1.2.5 Learning assessments for the sake of learning assessments?

Countries must demonstrate they have either a system in place to monitor learning outcomes, or a plan to develop one to receive a programme implementation grant from the GPE (GPE, 2019c). Consequently, learning assessment systems are expected to be an integral part of ESPs in the case of countries that are eligible for GPE funding. An important finding of our research is that ESPs do indeed strongly focus on the development of learning assessment systems, whereas less attention is paid to policies ensuring effective actual use of assessment data.

Sector plans typically acknowledge current weaknesses of learning assessment systems and the use of the data produced (see *Table 3*).

1. Overview of the use of learning assessment data in planning cycle phases

Table 3. Challenges identified in ESPs in learning assessment systems

Education sector plan	Challenges identified
Ghana Education Sector Plan (2018–2030)	— sample-based design of assessments — lack of 'national annual assessments at primary level' — lack of a 'comprehensive and annual assessment at certain key points of education' — absence of a 'comprehensive assessment policy to guide learning outcomes' (Ghana, 2019: 27)
Guinea ProDEG (2020–2030)	— weakness of standardized assessment and national examination systems — virtual absence of the use of assessment data generated for quality management
Namibia Education and Training Sector Improvement Plan (2005–2020)	— few mechanisms in primary and secondary education for the measurement of the system's performance against international benchmarks
Senegal PAQUET-EF (2013–2025)	— poor learning quality is attributed to the inadequacy of the learning assessment system — absence of a policy and unified mechanism — discontinuity of the national assessment (SNERS) — dependence on external funding — poor use of assessment results
The Gambia Education Sector Strategic Plan (2016–2030)	— inadequate validity and reliability of results received from regions — weak capacity to manage assessments in some areas/regions — poor quality of assessments

1. Overview of the use of learning assessment data in planning cycle phases

Education sector plan	Challenges identified
Zambia Education and Skills Sector Plan (2017–2021)	No challenges are explicitly mentioned, but strategies for improving the use of learning assessment data are specified: — expand the use of learning assessments as guided by the National Assessment Framework — align learning assessments with other monitoring systems in view of decentralization process

Source: Compiled by authors.

ESPs often include strategies to improve, institutionalize, and expand learning assessment systems and some call for the creation of assessment policies (e.g. the Guinea ProDEG). Other strategies focus on strengthening assessment systems through improved dissemination, results analysis, and regularity of assessments. Some ESPs plan for the expansion of learning assessment systems through either the creation of or participation in new assessments, or the increased scope of existing ones to include other types of schools or grades (e.g. secondary education). For example, the Guinean Ministry of Education (Ministère de l'Éducation nationale et de l'Alphabétisation, MENA) expressed its willingness to expand a national assessment to secondary education. Likewise, the Education Sector Medium-Term Plan (2008–2011) of The Gambia urged for the institutionalization of EGRA and the NAT in the madrasa schools by standardizing assessment procedures and translating materials into Arabic.

Although assessments are not new in most of the countries, systems are continuously willing to expand. The analysis of the discourse on learning assessment systems in ESPs provided by Uliana Furiv (unpublished) shows that this discourse has evolved over time, going from a wishful tone in the early 2000s

to a more action-oriented approach nowadays. Nevertheless, the ambitious strategies for developing assessment systems contrast with the current limited use of the existing data.

As demonstrated above, some phases of the planning cycle are more systematically informed by learning data than others, and the same is true for specific types of policies. The next subsection reports on related trends, as well as the uneven use of different types of assessments in policy-making.

1.3 Learning assessment data inform a narrow range of policies

Although learning assessment data can inform a broad range of policies, some trends in their use emerge from our analysis. These findings are in line with what was reported in a literature review on the use of learning assessment data conducted at the inception of this project (Raudonytė, 2019).

1.3.1 Types of policies informed by learning assessment data in the countries studied

Curriculum revision or development. Learning assessment data informed curriculum changes in all project countries. Many paid more attention to early grade literacy and numeracy, and modifications to the pedagogy used to teach them were introduced as a result of EGRA/EGMA assessments. For example, in The Gambia, learning data are used in the ongoing development of the early childhood education (ECE) curriculum. Similarly, in Zambia, NAS and EGRA results informed the extension of the curriculum to include national languages, while in Ghana TIMSS results encouraged changes in the science, technology, engineering,

and mathematics (STEM) curriculum to put more emphasis on these subjects.

Pre-service and in-service teacher training. Learning assessment data are often used to inform pre-service and in-service teacher training programmes. For instance, Namibia started the In-Service Education and Training (INSET) programme to improve the skills of currently unqualified and under-qualified teachers teaching at junior primary phase after NSAT and SEACMEQ tests revealed the significant impact teachers' qualifications have on student performance. In The Gambia, learning data also triggered a reform of pre-service and in-service teacher training based on EGRA results (Senghor, 2014).

Teaching methods. A number of programmes for the improvement of teaching methods have also been developed in the light of assessment results. For instance, in Zambia, low levels of reading in early grades led to the introduction of a catch-up programme, which promotes a targeted learning approach by regrouping pupils by their proficiency levels rather than by age. Findings from the assessment data supported early teaching in the mother tongue in Senegal and Zambia. In The Gambia, new teaching methods informed by EGRA focused on foundational reading skills, including pre-reading skills, oral reading, written comprehension, oral comprehension, and spelling.

Other policies were also mentioned in the interviews. In The Gambia, interviewees reported that the NAT results inform teacher posting, i.e. teachers with strong skills in a subject area are transferred to schools where there is poorer mastery of these subject areas. Moreover, EGRA served as a tool in the

development of a handbook on teaching early grade reading abilities. In addition, according to the Zambia Systems Approach for Better Education Results (SABER) report (World Bank, 2009), SEACMEQ results informed school resource allocation in Zambia. The development of a textbook policy was informed in Namibia. In Guinea, Senegal, and Zambia, learning assessment data are used to support initiatives for girls' education. In Namibia, an interviewee mentioned that SEACMEQ triggered a policy on educationally marginalized children.[12]

1.3.2 The comparative influence of different learning assessments

The types of policies influenced by learning assessment data are intrinsically linked to the assessments that inform them. In many instances, EGRA results had a more significant impact on policy formulation than any other assessment.

EGRA's development was motivated by the need to gather timely information on and for learning improvement in low-income countries (Dubeck and Gove, 2015). Its design is based on research on the development of reading skills. Since its creation in 2006 by RTI International (contracted by USAID), it has been implemented in 65 countries and adapted into 100 languages (Dubeck and Gove, 2015). The assessment has gained an increasing appeal, partly due to its adaptability to national contexts and its more focused sample size, which makes it 'smaller, quicker, and cheaper' (Wagner, 2011).

12 'The Ministry of Education, Arts and Culture (MoEAC) is developing and implementing policies to include more educationally marginalized children. Key among these are the Education Sector Policy for Orphans and Vulnerable Children (2013), and the Sector Policy on Inclusive Education (2013).' (UNICEF, n.d.:1)

1. Overview of the use of learning assessment data in planning cycle phases

It is important to note that EGRA is different from most LSAs. This assessment is oral, administered face-to-face and generally in schools. The assessments have been primarily designed to provide simple, low-cost measures of reading (UNESCO, 2019). Although EGRAs rely on a common methodological framework across countries, they are not intended to enable cross-country comparability (UNESCO, 2019). The results are sensitive to the characteristics of the language in which assessment is administered. Moreover, EGRA results may be representative at national or subnational levels, but this depends on the design of the assessment's implementation, which varies greatly from one country to another.

EGRA results were heavily referenced in ESAs of our project countries, and they influenced the design of multiple policies. Policies related to teacher training, pedagogical methodology, or curriculum have been developed following EGRA results, or according to the EGRA conceptual framework;[13] as described in the previous section, these are precisely the types of policies that have been most influenced by learning assessment data.

In the case of The Gambia, there is strong evidence in the literature (see, for example, World Bank, 2011) and some evidence in the country's recent Education Sector Strategic Plan (2016–2030) to suggest the prioritization of EGRA over the NAT data when informing education strategies, particularly instruction- and curriculum-related policies at

13 Dubeck and Gove (2015) underlined the fact that EGRA is not an instructional programme per se, but it can be used to inform programmes and interventions under the EGRA banner (e.g. an eponymous experimentation programme in Guinea). Hence the confusion between the assessment and reading programmes that may follow the same conceptual framework: 'As a formative assessment, teachers can either use EGRA in its entirety or select subtasks to monitor classroom progress, determine trends in performance, and adapt instruction to meet children's instructional needs' (Dubeck and Gove, 2015: 2).

primary level. EGRA and EGMA played an important role in informing policies, reviewing standards, promoting the use of national language in early grade instruction, as well as developing teaching methods, materials, and teacher training. In contrast, the NAT results have been used at decentralized level, especially by schools, for strategic management of resources, planning, and introduction of targeted interventions (e.g. remedial classes for struggling students). The NAT results have also been used to set baselines and targets in the Education Sector Policy for 2022 and 2030.

It is difficult to account for this heavier use of EGRA results in sector plans or sector analyses, but our results suggest some possible explanations. Firstly, this assessment is often an integral component of reading projects supported by development partners. In Senegal, the EGRA assessment is embedded in the *Lecture Pour Tous* programme sponsored by USAID; in Guinea, it is used as the impact evaluation for a new approach to teaching reading. The integration or close linking of EGRA with such major programmes has strengthened its impact and facilitated the translation of its findings into policy. Secondly, the mastery of reading skills in early grades has become an area of focus in all countries, which may enhance EGRA's relevance in the landscape of learning assessments. However, the assessment might also have contributed to this increased attention on the issue.

In comparison, our research did not find many examples of PISA-D, PASEC, and SEACMEQ data informing planning and policy-making. In addition, many research participants complained about their limited potential to have an impact in classrooms but also in planning activities at decentralized level, mainly because of limited data availability and

1. Overview of the use of learning assessment data in planning cycle phases

disaggregation. Addey et al. (2017) support this finding: 'Technical questions about how [international large-scale assessments] ILSAs can sensibly inform better teaching and learning practices are increasingly being voiced by those concerned that resources invested in ILSAs give something back to education at the classroom level. However, in the case of PISA it is not clear how ILSA data could sensibly inform changes at this level.' (Addey et al., 2017: 9) Nevertheless, international assessments have played a role in the introduction of several programmes. In Namibia, respondents reported that a policy on educationally marginalized children and a textbook policy was influenced by SEACMEQ results.

Our study did not allow for documenting the use of data from citizen-led assessments in an in-depth way. In our project countries, only Senegal implemented such an assessment (i.e. Jàngandoo). Its results informed some partner projects, including the Lecture pour Tous USAID programme, PADES (Programme d'Appui au Développement de l'Éducation au Sénégal) led by the French Development Agency, and some United Nations International Children's Emergency Fund (UNICEF) initiatives. Moreover, a Senegalese Development Plan (*Plan Sénégalais Émergent [2019–2023]*) also relied on some of its data.

In conclusion, national tests such as the NAT may be less used to inform specific programmes because they are better suited to serving the purpose of diagnosis, whereas EGRA/EGMA are often developed to respond to a specific type of intervention. In other words, the primary purpose of EGRA/EGMA may not be diagnosing the education system's performance as a whole, although in some cases they were

used for this purpose when implemented on a nationally representative sample. Consequently, the purpose of the assessment has an impact on the extent to which it can inform broader policy direction (diagnosis-focused assessments) or specific strategies (more focused assessments). International assessments are not easily linked to a context-based policy. They can be useful for comparing the performance of different countries or as a reference for what kinds of strategies may work in similar contexts. Based on this analysis, it is not possible to recommend one type of assessment over another, but it helps to see that different assessments serve some objectives better than others. It is important to consider these elements when choosing an assessment to implement.

Chapter 2. Lessons on the institutional setting for the use of learning assessment data

An established institutional setting can support more effective use of learning data. It can provide a framework for more purposefully designed and implemented assessments managed by well-organized government structures. This chapter analyses institutional settings that exist in the countries studied for the use of learning data, exploring their assessment policies, assessment modalities, and the institutional arrangements of assessment teams.

2.1 The patchwork of learning assessments

Numerous assessments available in countries often do not coordinate well with each other, for example because they evaluate the same grades. Assessments often have overlapping goals and do not provide data that could complement each other to generate new insights. A lack of coherence and complementarity among various assessments contributes to less effective use of learning data. This issue was prominent in several countries we studied. For instance, Ghana has numerous learning assessments, but these are disjointed and poorly coordinated, without an overarching guiding strategy for national assessments. Representatives of a regional education office noted: 'We all use data from different sources. That makes it difficult for us to achieve the same goal.' In Namibia, SEACMEQ and NSAT assess three consecutive grades (Grades 5, 6, and 7), which does not allow problems to be identified in early and higher grades. In addition, information for reporting on SDG indicator 4.1.1

2. Lessons on the institutional setting for the use of learning assessment data

is missing.[14] In Senegal, several assessments cover the same years (e.g. SNERS and PASEC both assess Grade 2) or are conducted at short intervals (e.g. PAQEEB and ESH assessments). Inconsistencies across assessments in the system in turn lead to the following issues that negatively affect the use of learning data:

- Scarce financial and human resources are dispersed across overlapping assessments, which reduces efficiency in the use of funds, and is particularly problematic in contexts where they remain low. For example, this issue is very salient in the Senegalese assessment system.
- Poorly coordinated assessments provide data that are not complementary and are therefore rarely analysed together. Doing so becomes even more difficult when the number of assessments is significant. In Senegal, learning data coming from at least seven different assessments are analysed in isolation without mechanisms to capitalize on the richness of different data sets.
- The differentiation between assessments and their goals becomes less noticeable; actors are less clear about the purpose of different assessments. It is illustrative that many of those interviewed were not very aware of differences between assessments and there has been systematic confusion around definitions of assessments and examinations. At decentralized levels in Guinea and Ghana, this confusion was clearly evident in the interviews.

One of the factors that both contributes to and arises from the problem examined above is the lack of a regulatory

14 Indicator 4.1.1: Proportion of children and young people (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex.

framework for learning assessment systems and the use of their data. A large body of literature emphasizes that the institutionalization of assessments is an important factor influencing the use of data; this can be improved if assessments are not seen as one-off isolated exercises and if they are integrated into regular ministry operations (Raudonytė, 2019; Tobin, Nugroho, and Lietz 2016; DFID, 2011).

The institutionalization or integration of assessments can refer to one of the following activities: '(a) mandate the establishment of an assessment agency, (b) outline the regular conduct of an assessment programme over time and (c) fund the activities of an assessment programme'[15] (Tobin, Nugroho, and Lietz, 2016: 590). The rest of the chapter examines the two first dimensions (assessment policies and assessment modalities, and assessment units [AUs]).

2.2 Assessment policies either non-existent or poorly disseminated

Studies emphasize the importance of clearly defining the purpose of all assessments that should help guide their design and implementation (Elks, 2016). The literature underlines that assessment policies have a strong potential to improve assessment systems and the use of their data, especially by providing a strategic vision for assessment systems:

> Learning assessment needs to be guided by legislation or policy in order to gain meaningful evidence at all levels of the education system, from national and subnational levels, to school and classroom level to the level of individual learners. That is, learning

15 The funding of assessments is addressed in *Chapter 3*.

2. Lessons on the institutional setting for the use of learning assessment data

> assessment needs to be *purposefully* designed, implemented, analysed and disseminated to provide the adequate evidence for the education policies and practices they aim to inform. (GPE and ACER, 2019: 5)

Assessment policies/frameworks[16] can serve the following objectives, among others:

- define the overall purpose/s of the assessment;
- define 'measurement priorities', i.e. specific statistical objectives addressed by the assessment;
- identify the primary users of the assessment data and clarify how their needs can be met;
- specify characteristics of the learning assessment, i.e. what it will cover in terms of content, skills, knowledge, and context;
- describe the kind of data an assessment will provide, and how that data are going to be reported;
- help the wider community interested in the assessment understand what it is about and what the assessment results mean (ACER-GEM; UIS, 2017).

A Zambian assessment framework is an interesting example covering most of these elements (see *Box 2*).

> An established institutional setting can support more effective use of learning data.

16 Assessment policy and framework can come in two separate documents or constitute one. Whereas a policy defines an assessment system in more general terms, a framework is likely to provide more technical aspects on assessments. However, in some instances, these terms can be used interchangeably.

2. Lessons on the institutional setting for the use of learning assessment data

BOX 2

Assessment policy in Zambia

Although not yet implemented, the National Learning Assessment Framework (NLAF) (Zambia, 2017) is an overarching document that defines the principles that apply to all student assessments in Zambia. It defines the intended use of learning assessment data, both in general and for specific actors. It outlines the responsibilities of actors, as well as existing types of assessments, and demonstrates how they articulate with each other. It provides consistent definitions of the key concepts linked to assessments. The document also provides guidance for the diagnostic use of assessment information to plan remedial action. Most importantly, it sets out conditions for effective decision-making using assessment results, such as timely availability of data.

Source: Adapted from Raudonytė and Bodin, 2021.

2.2.1 What was found in participating countries

In the countries studied, a proliferation in student assessments, strongly encouraged by external partners, was not accompanied by a subsequent timely development of related national policies. Consequently, assessments were often introduced with little consideration of how they would articulate with the existing ones, and to what extent they could, in combination, address national policy concerns.

> “ the institutionalization of assessments is an important factor influencing the use of data.

2. Lessons on the institutional setting for the use of learning assessment data

Four out of six countries in our study do not currently have an assessment policy in place although its development is planned in three of them (Guinea, Senegal, and Ghana). A 10-year strategic education plan in Guinea specifically plans for the development of an assessment policy, and preparatory activities for its development started in Senegal and Ghana. The Gambia and Zambia developed assessment policies a couple of years ago (see *Table 4*).

Table 4. Assessment policies in the countries studied

Country	Status of the assessment policy
Ghana	development foreseen
Guinea	development foreseen
Namibia	no policy
Senegal	development foreseen
The Gambia	policy in place but dissemination missing
Zambia	policy in place but not fully implemented

Source: Compiled by authors.

This does not, however, mean that there are no official documents providing guidance on assessment implementation. Elements are found dispersed across a number of documents (e.g. external programming files, circulars, *arrêtés*, assessment conceptual frameworks) that do not constitute a strategic vision for assessment systems.

It is difficult to explain this lack of or belated development of assessment policies. Our study advances one possible explanation: one respondent in Ghana noted that perhaps because assessments have so far been largely funded by

external sources and started as external projects, they were not part of ministerial structures and therefore policies. Nevertheless, this question could not be addressed in depth in our study.

2.2.2 Actors' awareness about existing regulatory frameworks

We also examined to what extent actors in the system were familiar with documents regulating learning assessment systems and the intended use of learning data. Participants would often be able to define the main goals of assessments, albeit in quite general terms. However, it was much more difficult for many of them to speak about legal documents that regulate assessment systems, as in many cases specific documents do not exist and only a range of dispersed elements are available. Actors' knowledge of the regulatory framework depends on their closeness to the assessment system and whether they were involved in consultations to develop it. Decentralized-level officials had greater difficulties in referring to regulatory documents than their counterparts at central level.

For example, in Ghana, officials at central level have varying degrees of knowledge regarding the existing set of assessment documents, depending on their closeness to the assessment system: some are aware of the development of the assessment policy and could provide some details on it, whereas others only mentioned external programming documents. However, officials at decentralized level (i.e. regional and district education officers) were not aware of the regulatory documents on learning assessments; they are currently more familiar with examinations.

2. Lessons on the institutional setting for the use of learning assessment data

2.2.3 Preconditions for a more effective assessment policy

Institutionalization of assessments, through the development of an assessment policy, is an important step towards improving the use of learning data. Nevertheless, it will not be sufficient if certain conditions are not in place. Inclusive policy development processes, as well as wide dissemination, are key to ensuring its effectiveness. The assessment policy in The Gambia lacked wide dissemination and therefore wider application. While it provided guidance to actors directly involved with assessments, it was much less relevant to officials at decentralized level, who are further removed from assessments' design and implementation. In Zambia, the NLAF development was restricted to a limited number of stakeholders, which led to low national ownership of the document, which in turn weakened its implementation.

The Australian Council for Educational Research (ACER) 'Principles of Good Practices in Learning Assessment' underlines the importance of inclusive development of assessment frameworks to improve the use of learning data:

> Consult stakeholders. ... This could include presenting it to a steering committee in which various stakeholders are represented ... and/or aiming for a larger audience of policy-makers and learning domain experts. By ensuring that various stakeholders have had the chance to comment on the details of what is assessed, it is more likely that results will be accepted and used in improving outcomes for learners. However, responsibility for finalizing the framework should rest with a combined team of experts (expert committee) and test developers. (ACER-GEM; UIS. 2017: 17)

Alignment of an assessment policy with other key documents, such as the curriculum, teacher framework, or an ESP, is another important precondition for an effective institutional

framework for linking national assessments with other aspects of the education system (e.g. curriculum, teacher education, school capacity-building, and measures to address inequalities (DFID, 2011)). Zambia is an interesting example, as the NLAF articulates well with two other key policy documents, i.e. the Zambia Education Curriculum Framework and the Teacher Curriculum Implementation Guide. Moreover, based on these documents, teacher training modules on assessments have been developed recently but are not yet in place. In The Gambia, the content of an Assessment Policy is closely linked to an Education Sector Policy (2016–2030) and an Education Strategic Sector Plan (2016–2030). This has helped both countries to insist on assessment-related strategic activities across education sector policies.

> Inclusive policy development processes, as well as wide dissemination, are key to ensuring its effectiveness.

2.3 Challenges with current assessment designs

When introducing assessment programmes, countries face a number of decisions with regard to their design and these should be made in the light of MoE information needs, as well as the officially defined purposes of assessments. MoE information needs differ per administrative level or unit (e.g. curriculum, teacher training) and need to be systematically identified and aligned with the purposes of assessments. The choice of grades and subjects assessed or sample size will influence the information that is produced, and this may have different implications for the use of assessment data, as is detailed further in this section.

2. Lessons on the institutional setting for the use of learning assessment data

Current assessment designs do not always respond to stakeholders' information needs, which are often not clearly identified. This constitutes an important barrier to the use of learning data. Consequently, our research demonstrates that some countries might benefit from reconsidering sampling methods, the population assessed, as well as the frequency of assessments. However, potential changes might entail significant consequences in terms of costs, which should be carefully evaluated beforehand.

2.3.1 What sampling methods to use?

Our research shows that in the countries of our study, there is often a government preference for census-based assessments, which echoes a rise in the share of census-based assessments that started in the early 2000s (Verger, Parcerisa, and Fontdevila, 2019). While The Gambia, Namibia, and Senegal have already adopted census-based assessments, in Ghana, the MoE has initiated a shift in the NEA design: '[t]he MoE is also planning to introduce a regularly implemented low-cost national assessment that can be administered to every pupil in grades 2, 4, 6 and 8, to be used to provide feedback and better instruction support' (World Bank, 2019b: 13). Multiple interviewees explained that because the NEA sample is not representative at district level, it does not allow challenges at this administrative level and in schools to be identified and addressed. Moreover, according to a retired official from Ghana Education Services, teachers have been excluding the lowest-performing students from the NEA sample, which constitutes another argument in favour of a census-based assessment. The future National Standardized Assessment Test will therefore be developed to provide information at school level, thereby allowing more

targeted interventions. Similarly, in Guinea, one representative of the Strategic Development Office (Bureau Stratégique de Développement [BSD]) explained that data could be more useful, and have better ownership, at the Ministry of National Education and Literacy if the national assessment covered all Grade 4 pupils.

Sample- and census-based assessments come with their respective advantages and disadvantages. Drawing on our evidence and the available literature, we present the potential implications of each choice.

Advantages of census-based assessments

An assessment can report on school and individual performance, thereby providing diagnostic data to each school and informing decisions on resource allocation in individual schools (DFID, 2011). This might be of particular relevance in countries with decentralized governance, and where subnational authorities have sufficient autonomy to plan interventions, as is the case with school improvement plans in Senegal or The Gambia.

Census-based assessments can inform parents and communities about individual school or student performance. (In The Gambia, learning assessment data are disseminated through scorecards at community level, showing school performance in the NAT and GABECE compared to the available resources and the performance of other schools in the region/country.) They are more appropriate for accountability purposes 'since they allow more performative pressure to be placed on all the schools' (Verger, Parcerisa, and Fontdevila 2019: 6). And they are also more equitable, as all students of a given level are assessed and disparities between schools can be analysed and highlighted.

Source: Compiled by authors.

2. Lessons on the institutional setting for the use of learning assessment data

Disadvantages of census-based assessments

Census-based assessments are more resource-intensive. As these assessments evaluate all students at a given grade, their cost is higher than that of sample-based assessments. Greater logistical, capacity, and budgetary challenges also emerge in test administration (especially in large countries, with long distances to travel and hard-to-access schools). In addition, more staff are needed to supervise test administration.

Although national LSAs do not have high stakes, when they are census-based, there is a risk that their stakes unintentionally increase and learning data are misused. For instance, in The Gambia, some schools have used the NAT results for student promotion purposes, as well as for the evaluation of teacher performance, fostering competition among them. These unintended consequences may encourage actors to distort assessment results (Senghor, 2014).

Source: Compiled by authors.

Advantages of sample-based assessments

As fewer students and schools are assessed, operational costs, and costs related to data treatment, are lower. This can make a significant difference in large countries (DFID, 2011). The amount of time dedicated to data treatment and analysis should also be lower (DFID, 2011).

Supervision of data preparation and fieldwork takes up fewer logistics, fewer staff, and less time, thereby ensuring higher data quality (DFID, 2011).

Source: Compiled by authors.

2. Lessons on the institutional setting for the use of learning assessment data

Disadvantages of sample-based assessments

One major disadvantage of sample-based assessments is that they cannot inform the work of specific schools and subnational stakeholders (depending on the level at which the sample is representative). For this reason, some subnational actors often question the relevance of sample-based assessment data for their work. These unaddressed needs may ultimately lead ministries to consider census-based assessments.

Drawing representative samples can be highly technical and require advanced statistical skills that are not always available in countries.

In some countries, teachers have been excluding the worst students from the assessment sample, which distorts overall assessment results. This is a reason why some MoEs (i.e. The Gambia, Ghana) are in favour of census-based assessments.

Source: Compiled by authors.

As *Table 5* shows, a census-based design seems appealing as it multiplies possibilities for the use of assessment data and responds to the information needs of a higher number of actors, especially officials and schools at subnational level. Nevertheless, for these advantages to materialize, some preconditions are needed. For assessment data to be used effectively, the introduction of a census-based assessment should go in hand with improvements in data dissemination, capacity development, and collaboration mechanisms to ensure data are available promptly at all levels. In The Gambia, although a census-based national assessment provides information on each school, it was initially difficult to use the NAT results to inform school improvement plans, due to head teachers' low capacities.

2. Lessons on the institutional setting for the use of learning assessment data

On the other hand, even though a sample-based assessment does not provide information on individual schools, it can, depending on the sample's representativeness, provide information on different population groups. If there are particular sub-populations of interest, their individual sample size needs to be large for valid statistical comparisons (ACER-GEM; UIS, 2017). Therefore, before envisaging census-based assessments, an alternative solution could be to increase the sample size, although this will still not provide information on individual schools and learners. Sample-based assessments are a valuable source of information at system level that can be produced at lower costs, and their advantages should not be overlooked.

It is not possible to determine which design has led to more effective use of assessment data. The choice from among these options should be guided by the goals of the assessment. However, Tobin, Nugroho, and Lietz (2016) explored two systematic reviews that examined evidence on the link between LSAs and education policy in economically developing countries and countries of the Asia-Pacific region. They found that sample-based assessments more frequently inform policy-making than census-based ones, but nuanced this by saying that '[t]his is likely to reflect financial and logistical considerations associated with the assessment, rather than any relationship to policy goals and uses, given the costly nature of census-based assessments' (Tobin, Nugroho, and Lietz, 2016: 582).

2.3.2 What target population to assess?

Choosing which grades to assess is an important decision. Countries can adopt different approaches when choosing

2. Lessons on the institutional setting for the use of learning assessment data

the level(s) to assess, but SDG indicator 4.1.1 provides clear guidance, as they have to report on the proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex. The SDG 4 Data Digest report (2018) noted that cross-national assessments are naturally at the forefront when reporting on SDG indicator 4.1.1, since they are designed for comparability purposes.

As shown in *Table 5*, few standardized assessments cover secondary education. Priorities placed on addressing low mastery of foundational skills (which needs to be identified at primary level) partly explain the prevalence of primary grades in national assessments. However, such a concentration of assessments on the same student population may hamper the monitoring of education quality throughout schooling and the possibility for longitudinal linking.

In Senegal, the PAQUET 2018–2030 calls for the establishment of a national system for monitoring and evaluating learning in secondary education. Currently, the Standardized and Harmonized Assessments (ESH) conducted at decentralized level target lower secondary grades, but they have not been institutionalized yet, and their implementation mostly depends on the initiatives of School Inspection Offices. In Guinea, the ProDEG plans for a national LSA at secondary level, but its feasibility is largely constrained by scarce capacities and financial resources.

Although the NAT also evaluates Grade 8 students in The Gambia, interviewees expressed concerns about the lack of interventions in the upper grades. To date, the impact of the NAT in Grade 8 remains largely unrecorded and unacknowledged.

2. Lessons on the institutional setting for the use of learning assessment data

Table 5. Reporting on SDG indicator 4.1.1: grades covered by learning assessments in the project countries

	Grades 2/3	End of primary education	End of lower secondary
Ghana	EGRA/EGMA		
Guinea	EGRA Évaluation Nationale des Acquis	PASEC	
The Gambia	NAT		NAT
Namibia		SEACMEQ	
Senegal	EGRA/EGMA SNERS PASEC	PASEC	PISA-D
Zambia	EGRA/EGMA	SEACMEQ	PISA-D NAS

Source: Compiled by authors.

In addition, as argued earlier, overlaps of assessments in some grades indicate that the choice of grades to assess is rarely made in a strategic way to ensure complementarity and coherence of the information produced. Even in countries with multiple assessments, learning assessment systems do not always allow for compliance with the SDG 4.1.1 benchmark.

Little evidence allows us to identify the rationale that motivates the choice of grades in national assessment designs. In Guinea, the *Rapport Bilan du PSE 2008–2014* indicates that, following the results of the 2007 Grade 4 *Évaluation Nationale des Acquis*, there has been a consensus that Grade 4 is a key point in the development of foundational skills in reading, writing, computing, and sciences. This justified the annual implementation of the Grade 4

assessment. Grades 2 and 6 are alternately evaluated every three years. In countries such as Namibia and The Gambia, the introduction of national learning assessments coincided with the phase-out of examinations in certain grades. In Namibia, the NSAT replaced examinations in Grade 7, while in The Gambia, the NAT was introduced following the phase-out of the Primary School Leaving Certificate Examination in Grade 6, though the former does not assess pupils of the same grade (The Gambia, 2015). In Senegal, however, grades assessed by SNERS have been constantly changing from one cycle to another. Nevertheless, some stability has been observed in the latest rounds of SNERS' implementation, as it has covered Grades 2 and 4.

Another important principle to consider in LSA design is its inclusiveness. This means to 'design assessments to be relevant for as many members of the target population as possible' (ACER-GEM; UIS, 2017: 8). As regards the use of assessment data, it means that policies can better address the needs of a potentially underperforming or under-resourced part of the student population. Nevertheless, our research shows that inclusiveness remains an issue in certain national assessments.

In Ghana, children with disabilities do not take national EGRA/EGMA assessments. Likewise, students in Special Schools (i.e. schools for children with visual impairments, hearing impairments, and severe learning needs) do not participate in the NEA (World Bank, 2013); moreover, private school presence in the assessment is limited. According to a non-governmental organization (NGO) representative, the government is primarily interested in public schools and it does not intentionally aim to reach private schools, even

though 20–25 per cent of pupils are registered at private schools. Similarly, EGRA and EGMA collect some data on private schools, but this is not systematic or intentional. In Senegal, the SNERS excludes schools where instruction is provided in Arabic. By contrast, in The Gambia, inclusiveness is one of the strengths of the learning assessment system. According to the 2017 NAT report (The Gambia, 2017), all government, grant-aided, and almost all private schools, participate in the NAT. Still, Islamic schools (madrasas) do not take part in the NAT yet and plans to incorporate them have been thwarted by a lack of financial and human resources.

2.3.3 At what frequency?

According to the Department for International Development (DFID, 2011), decisions on assessment frequency depend on two major factors: the use of assessments results and related costs. As achievement levels change very slowly, a three- or four-year time lag can be adequate for sample-based assessments to report on learning outcomes, while a shorter time period can be justified by particular circumstances (e.g. a recent reform) that require obtaining information more frequently (DFID, 2011). In the project countries, assessments' frequency is largely determined by the available resources. This can explain why they were not always conducted following the frequency initially decided upon and why there have sometimes been long periods of interruption.

With various assessments in place and many assessment features, Senegal epitomizes the debate on assessment frequency and its implications for the use of learning assessment data. When it comes to the PASEC assessment, its representative argued that a five-year cycle allows

2. Lessons on the institutional setting for the use of learning assessment data

the effects of the recommendations implemented to be perceived. However, other stakeholders at the Ministry of National Education and international partners felt that this cycle length was not consistent with the sector's dynamics, and that the data are not collected frequently enough. Regarding SNERS, ministry officials expressed a preference for a more frequent assessment; it has been irregular until now, often conducted every two to three years. One representative from the Direction de la Planification et de la Réforme de l'Éducation (Education Planning and Reform Department [DPRE]) explained that they wanted to conduct SNERS on an annual basis. The PAQUET 2018–2030 confirms dissatisfaction with the current frequency of SNERS, which largely depends on external resources: 'In addition, the frequency of the SNERS, which is dependent on the availability of external financial resources, is regrettable; moreover, its results are poorly exploited.' (Senegal, 2013: 105) Nevertheless, other interviewees reported that the tight annual periodicity of decentralized assessments does not allow stakeholders to fully engage in the data analysis and exploit the results. A similar issue was reported in Zambia where, although the NAS is conducted every two to three years, multiple interviewees explained that more time was needed to capitalize on assessment results and implement appropriate interventions.

It is noteworthy that the frequency of assessment implementation needs to be consistent with dissemination timeliness. Given current national challenges encountered when delivering assessment results in a timely manner, it might be less relevant to implement assessments at annual frequencies. For instance, in The Gambia, where the NAT is administered annually, participants complained about the

2. Lessons on the institutional setting for the use of learning assessment data

time it takes to release the NAT data, which then delays the entire chain of data analysis and use at schools. As observed in Namibia and The Gambia, a further risk is that annual assessments are perceived as a 'mechanical exercise' that fails to initiate critical reflections on learning assessment data. A university professor in The Gambia observed that 'people get used to it and they might miss the point and essence'.

On the other hand, long intervals between assessment cycles, coupled with lagging dissemination, can also create challenges for the use of assessment data. This issue was particularly salient with the regional SEACMEQ assessment, since many participants deemed its frequency and the timeliness of its results delivery (sometimes years after data collection) unsatisfactory. A planning and development (PAD) representative in Namibia explained that, by the time SEACMEQ results were out, 'it started losing momentum'.

Rationales for conducting annual LSAs can be motivated by the need to monitor students' performance more closely over time, but assessment characteristics then need to ensure data are comparable over time. In The Gambia, according to a World Bank education sector public expenditure review (2017), the current instruments (including the NAT) are not appropriate for assessing learning outcomes over time, though a World Bank representative reported that efforts are made to introduce more comparability. This is an important parameter to consider, as changes in the assessment system can introduce distortions in comparability and therefore the monitoring of learning outcomes over time (e.g. in Ghana).

> " the frequency of assessment implementation needs to be consistent with dissemination timeliness.

2. Lessons on the institutional setting for the use of learning assessment data

The changes envisaged in assessment design have important implications in terms of cost but also in terms of the use of the data. Consequently, before any change in assessment design, it is necessary for ministries to consider potential repercussions and to critically evaluate the likelihood that such a shift improves the use of learning assessment data. More attention should also be paid to the effectiveness of the current learning assessment designs, as noted by an Examinations Council of Zambia report:

> This situation is typical of most national assessments; because the assessments are typically the evaluation metrics for other interventions, there is insufficient attention paid to whether the assessments themselves are successful. (Zambia, 2019a: 15)

2.4 Lessons on institutional settings for assessment units

Institutional arrangements for those responsible for assessments are key, as 'clear governance structures, institutional arrangements and accountability mechanisms are required to ensure national policies for learning assessment are implemented effectively' (GPE and ACER, 2019: 5). Different institutional settings are possible, and they can be classified along two main dimensions:

1) the closeness of the assessment team to the MoE (i.e. its functional autonomy) as categorized by UNESCO and UIS (UIS, 2018b) (see *Table 6*);
2) the level of concentration of different assessments and related activities (i.e. responsibilities related to assessments are concentrated in one unit/institution or they are dispersed among different units/institutions, as emerged in the research).

2. Lessons on the institutional setting for the use of learning assessment data

Table 6. Institutional settings for assessment management, in terms of closeness of teams to the MoE, advantages and disadvantages

Setting	Description	Advantages	Disadvantages
Unit within the ministry or department of education	Assessment team is institutionally located in the MoE as one of its units/ departments	Facilitated coordination between curriculum, pedagogy, and assessment teams, which in turn facilitates alignment of these areas in the education system and the use of assessment results; however, other challenges might hinder this collaboration	Risk of assessment teams being more vulnerable to political interference (e.g. not publishing or altering poor results)
Semi-autonomous public institution	Institutions often have their own budget and are accountable to the minister of education or to congress (e.g. national institutes of statistics, research centres, or quality assurance agencies), in addition to their regular tasks of leading national, regional, or international assessments	More independent from political or collegial influence	Risk of a lack of coordination and misalignment of the assessment with other policies of the education system

2. Lessons on the institutional setting for the use of learning assessment data

Setting	Description	Advantages	Disadvantages
Examination board or unit	Institution in charge of examinations for certification and/or selection, which in addition to its regular tasks leads national, regional, or international assessments	Institutional capacity and expertise of the examination team mobilized for other assessments	Risk of overwhelming an institution that already has a clear mandate
University, NGO, or equivalent	External entities take over assessment activities (entirely or partially)	Mitigates a lack of technical expertise available in national assessment teams	Risk of limiting national ownership and leadership of the assessment system and interrupting a national cycle of learning data management

Source: Adapted from UIS, 2018b, and enriched by authors based on study results.

2.4.1 What was found in the selected countries, and what lessons can be drawn?

In the countries studied, different combinations of the two dimensions exist. Most often, an assessment team is located within the MoE (e.g. The Gambia, Guinea, Ghana); in addition, while in three countries (Namibia, Zambia, The Gambia) assessment-related activities are dispersed across units, in the remaining ones they are concentrated in one institution, as can be seen in *Table 7* and *Figure 4*.

2. Lessons on the institutional setting for the use of learning assessment data

Table 7. Institutional settings in project countries

	Institutional setting (two dimensions)	Comments
The Gambia	— Responsibilities linked to assessments are dispersed across different institutions. — The main unit in charge of a national LSA is located within the ministry.	The Assessment Unit (AU), West African Examinations Council, Planning Policy Analysis Research and Budgeting Directorate (PPARBD), Curriculum Research Evaluation and Development Directorate (CREDD), and Standards and Quality Assurance Directorate (SQAD) are involved in the management of NAT, with the AU playing a leading/coordination role; responsibility for EGRA assessment has been shifting.
Ghana	— Responsibilities linked to assessments are concentrated in one institution. — Unit in charge of a national LSA is located within the ministry.	National Education Assessment Unit (NEAU), under the National Council for Curriculum and Assessment (NaCCA), is in charge of the NEA and EGRA/EGMA.
Guinea	— Responsibilities linked to assessments are concentrated in one institution. — Unit in charge of a national LSA is located within the ministry.	Service National de l'Evaluation du Système Educatif (SNESE), under Institut National de Recherche et d'Action Pédagogique (INRAP), is in charge of the Évaluation Nationale des Acquis and INRAP is in charge of EGRA.

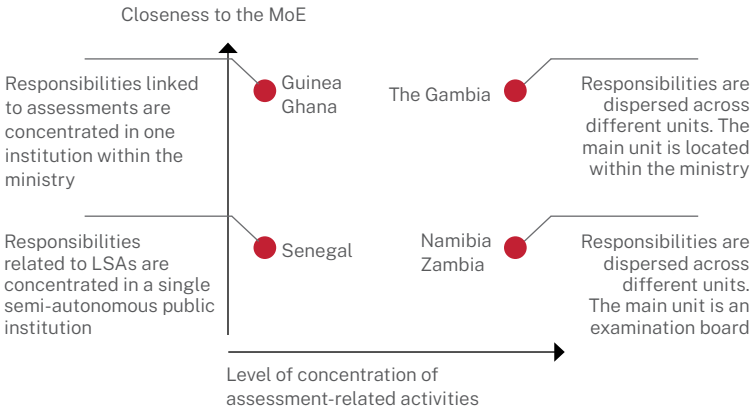
2. Lessons on the institutional setting for the use of learning assessment data

	Institutional setting (two dimensions)	Comments
Namibia	— Responsibilities linked to assessments are dispersed across different institutions. — Unit in charge of a national LSA is an examination board.	Each of the three key directorates is in charge of one learning assessment: a PAD unit leads SEACMEQ studies, a Directorate of National Examinations and Assessment (DNEA) is working on NSATs, and a National Institute for Educational Development (NIED) is responsible for EGRA. DNEA is also in charge of examinations.
Senegal	— Responsibilities linked to assessments are concentrated in one institution. — It is a semi-autonomous public institution.	Learning assessments, in almost all their aspects, are essentially the responsibility of Institut National d'Etude et d'Action pour le Développement de l'Éducation (INEADE); even for decentralized assessments, it provides technical assistance at different stages.
Zambia	— Responsibilities linked to assessments are dispersed across different institutions. — Unit in charge of a national LSA is an examination board.	Over the years, the Examinations Council of Zambia (ECZ) took more autonomy and became the leading institution for NAS and EGRA/EGMA; it is also in charge of examinations; the Department of Planning and Information at the MoE is in charge of coordinating SEACMEQ.

Source: Compiled by authors.

2. Lessons on the institutional setting for the use of learning assessment data

Figure 4. Visual representation of institutional settings for assessment management



Source: Developed by authors.

Bringing assessments institutionally closer to units making decisions in different education areas, especially curriculum and planning, might increase the use of assessment data for those purposes, but it disperses technical expertise on assessments as different assessment teams often do not collaborate in an effective way. Units in charge of the curriculum often host EGRA/EGMA assessments (i.e. as in Namibia, Guinea, and The Gambia). This strategic choice aims to strengthen the link between assessments and the curriculum. It encouraged the use of EGRA results to introduce changes in the literacy curriculum in the three countries. Nevertheless, curriculum departments do not necessarily have the same expertise as AUs. This was especially true in Guinea when INRAP had to externalize many assessment-related activities. In addition, this often results in poor communication between teams in charge of EGRA/EGMA and

2. Lessons on the institutional setting for the use of learning assessment data

a national or a regional assessment, and those in charge of other assessments, as was the case in Guinea and Namibia. In addition, planning units are the ones in charge of SEACMEQ assessments in Namibia and Zambia. This was encouraged by SEACMEQ coordinators, to draw a more direct link between assessment results and planning as well as policy-making. However, in both countries, this led to poor collaboration between specialists in charge of different assessments.

From an institutional perspective, being placed under the same umbrella organization, or more broadly an MoE, does not necessarily lead to increased collaboration between units that have responsibilities in assessments or those in charge of the curriculum, teacher training, pedagogy, and planning. Communication issues persist in most of the project countries. For example, in Ghana, the NEAU is placed within NaCCA, which has overall responsibility for the national curriculum and student assessments during the first and second education cycles (please refer to *Chapter 4,* which analyses national actors' collaboration in depth).

So which institutional setting is more conducive to improved use of learning data? Is it better to have one institute, such as INEADE in Senegal, or many organizations sharing the work, as in The Gambia, to manage assessment cycles? Is it preferable for an assessment team to be closer to the MoE or, on the contrary, is it preferable to transfer assessment activities to an independent body? Although there are a number of associated advantages and disadvantages to consider when choosing from among the available modalities, as previously examined, there is no one model that has an apparent strong overall advantage in all settings.

2. Lessons on the institutional setting for the use of learning assessment data

Existing literature underlines that 'well-established, legislated assessment agencies that have developed a high degree of regard with the public and education stakeholders are more likely to be protected from government regime change and assessment results more likely to inform education policy' (Tobin, Nugroho, and Lietz, 2016: 590). Indeed, the credibility of the assessment team in terms of its capacities and impartiality are key, if actors are to take assessment results seriously. Our research data confirm this; as these conditions were missing in Guinea, the trust in the assessment data and their use in policy were weakened. In addition, independently of the institutional modality, 'the institution leading the assessment should be accountable to a clearly recognizable body (e.g. the minister of education, congress or a national education commission) that is itself accountable' (UIS, 2018b: 162), to further increase the credibility of the assessment team.

To conclude, in most of the countries studied, there is often a complex system of assessments combined with a low degree of their institutionalization, which negatively affects the use of learning data. Assessment policies are either under development or poorly disseminated; national actors are therefore unaware of the existing regulatory framework. Assessment teams do not always benefit from high esteem and credibility in the eyes of other actors, which is an important factor that negatively affects the use of learning data. However, strong policies and institutions alone will not ensure effective production and use of learning data; other factors are key to in-depth understanding of the dynamics behind the use of learning data, as demonstrated in the upcoming chapters.

Chapter 3. Dynamics between barriers and conducive conditions for the use of learning assessment data

This chapter explores the first set of barriers and conducive conditions for the use of learning assessment data, focusing on their interactions. Although the available literature has explored some of them, the chapter focuses on information specific to project countries and the region. It will explore the following factors that have been found to shape how student assessment results are used:

— assessment reports;
— means of dissemination;
— links between the EMIS and learning assessment data;
— capacities of national officials to use and produce data;
— issues related to the financing of LSAs.

3.1 Assessment reports are a critical vehicle for the use of learning assessment data

Assessment reports are a very important means of dissemination, especially as countries often do not develop an extensive range of other dissemination products. Just as with other dissemination means, when easily and broadly accessible, they 'increase the likelihood that results will be considered and used by a variety of stakeholders in decisions about education policy and practice' (ACER-GEM; UIS, 2017: 44). Their content and presentation are therefore pivotal to the interpretation of data and their policy implications.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

This section looks at how assessment reports analyse and present learning assessment data in the countries studied. It also draws on end users' opinions to examine whether assessment findings convey policy messages effectively. The analysis of their content reveals that national assessment reports often suffer from shortcomings in terms of the robustness and thoroughness of data analysis. In addition, relevant lines of action are not always clearly identified. It is important to mention that regional and international assessment reports rarely suffer from these issues. Our research also suggests that, in addition to relevant content, it is key to ensure assessment reports are user-friendly and formatted for easy understanding. This is even more important in countries facing capacity challenges.

3.1.1 Analysis of learning assessment data in national assessment reports often limited

While assessment reports aim to provide an accurate picture of education system performance, as well as relevant recommendations to educational planners and policy-makers (ACER-GEM; UIS, 2017), the analysis of learning assessment data provided in national assessment reports often does not offer all the elements necessary to achieve this.

Learning results are often expressed by mean scores. Learning outcomes are generally reported as students' average scores, which can then be disaggregated per subject, gender, region, and location (urban or rural), type of school, etc. For instance, the 2013 NEA technical report in Ghana provides this type of analysis. While it provides a satisfactory overview of learning outcomes, the pervasive usage of average scores might hide existing disparities. Wagner, Wolf, and Boruch (2018) argue

that while analysing national learning performance through mean scores aims to raise national learning levels, it leaves out struggling learners. Thus, greater attention to learning distribution across the country is particularly needed, especially in countries with significant inequalities. A more detailed analysis of the distribution of student scores can help gain a better understanding of discrepancies in performance and facilitate the design of targeted policies; assessment results can identify population groups that are not learning and that deserve special attention, as well as those that are performing exceptionally. Another risk is that considering mean scores alone, when these are low, may give the impression that scoring above average is a good performance in relative terms, while it does not reveal much about the absolute educational attainment to achieve. Therefore, an analysis of mean scores has to be interpreted against the backdrop of learning achievement targets to gain meaningful insights.

Distribution of students' performance as per proficiency levels. Some assessment reports present student scores according to proficiency levels. However, these are not always clearly defined or, if defined, sometimes over-simplify the presentation of students' performance, omitting important information on learning differences. For example, the NEA (2016) notes that: 'the NEA's criterion for "proficiency" … was established in 2005 with the first NEA, and is based on answering just over half of the items correctly (i.e. ≤ 55 per cent) and thus does not effectively identify pupils who have a full grasp of the curriculum – that is, who are truly proficient in the subject area' (Ghana, 2016: 7). Such a definition of proficiency levels does not allow for differentiation between the best and worst students on different sides of the proficiency threshold. More importantly, proficiency levels

often inform ESAs and ESPs, and it is essential to define them appropriately. Outdated or misaligned proficiency indicators might lead to underestimating or overestimating students' knowledge and neglecting groups of students with latent learning gaps.

Analysing factors influencing learning. Policy-makers and planners are not only interested in knowing *what students learn*, but also *why some of them learn better and faster than others*. Indeed, many policy questions concern the relationship between learning outcomes and explanatory factors. This analysis can help understand to what extent and in what way student scores are influenced by a variety of demographic, in-school, and extra-school factors. In doing so, it can draw on the information provided by the background questionnaires administered along with assessment tests. An alternative solution is to link learning assessment data with the EMIS or surveys (e.g. household surveys) providing similar information. For instance, in Ghana, researchers linked student results to school characteristics available in the EMIS for the NEA 2011. Assessment data can be analysed in a number of ways, but regression analysis is the most commonly used technique (ACER-GEM; UIS, 2017).

Nevertheless, we identified this type of analysis as one of the most challenging tasks in the development of national assessment reports. In Ghana, Guinea, The Gambia, and Senegal, interviewees expressed frustration, as they were not able to get a deeper understanding of the causes of low student performance. Our research shows that assessment reports either do not contain an analysis of key factors influencing learning (e.g. NEA (2016), *Évaluation Nationale des Acquis* (2013; 2017), Grade 5 NAS (2014), and NAT (2016)), or

they present a limited analysis. For instance, in the SNERS 2016 report, the analysis of influencing factors relies on sets of variables (e.g. 'variables related to pupil characteristics', 'variables related to school heads and schools') rather than individual ones, to run a regression aiming to explain variance in student scores. This limits the extent to which inference can be drawn, as each set may contain a variety of variables that would deserve to be evaluated individually. It is indeed critical to know the weight and type of influence (i.e. positive or negative) of each factor on learning. Another example is the NEA 2013 report, which conducts a multivariate analysis that only includes demographic variables (i.e. variables on teaching and learning environment are missing). These issues with national assessment reports are, to a certain extent, related to a lack of capacities within national teams to conduct such analysis (see more on capacities in *Section 3.4*).

Overall, for assessment reports to spur action, policy-makers and planners need to have a comprehensive view of influencing factors, so that they can inform solutions for those that fall within the scope of educational planning and policy-making (e.g. teacher training, management of schools). However, it is also important to cross-check evidence coming from other sources, as assessment results alone cannot meaningfully inform policy action. It is crucial to keep in mind that they often only provide correlations, and that these need to be interpreted with caution in a broader country context and with other sources of information.

The robustness of the analysis signals the overall reliability and validity of assessment data, which account for the credibility of recommendations. The specification of certain elements that are routinely reported, along with regression tables (i.e.

standard errors, confidence interval, statistical significance, etc.) or other statistical tests conducted to analyse student performance can increase the robustness of the analysis in the eyes of statistically informed readers. However, this is not done systematically in national assessment reports. For instance, the SNERS 2016 report does not include regression tables. In addition, the Ghana report for EGRA/EGMA 2015 only lists variables that emerged as significant from the regression, without providing the parameters mentioned above. Concretely, this could have enabled an expert audience to evaluate results critically. It could also avoid over-simplistic or misleading interpretations by users. Gardner (1989) listed the omission of measurement errors as a factor conducive to test misuses. In addition, it is equally critical to know, out of the myriad of variables that were considered, what influences learning the most and what has a weaker relationship, or no relationship, with learning quality as per learning data analysis. The fact that an input emerges as non-significant may have important implications in terms of planning as well, because it can lead to revising priorities and diverting efforts from areas that could have been erroneously thought of as important for learning or that have diminished in importance over time.

3.1.2 National assessment reports do not provide enough inputs for policy-making

Similarly to the design of an assessment programme, assessment reports should be prepared keeping in mind the information needs of the final users, which may vary significantly (ACER-GEM; UIS, 2017). Our evidence suggests that recommendations provided based on analysis of national assessment results often fail to produce appropriate policy-making inputs.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

Firstly, in most of the national assessment reports analysed, recommendations do not offer specific insights for policy-making. They often remain too vague to provide guidance and need to be sharpened for a more concrete line of action. An example from the NEA 2011 report is illustrative:

> Thus, it is recommended that the focus in Ghana primary education now and in the years to come, be enhancing instruction in the early primary grades to ensure that these pupils have the foundational skills needed to succeed in school. (USAID, 2016: ix)

Sharpening policy recommendations would provide clearer guidance for planners and policy-makers. In addition, the manual of good practice on learning assessment suggests that 'implications or recommendations which have been discussed with stakeholders who have in-depth knowledge of the sector are likely to be more robust' (ACER-GEM; UIS, 2017: 45).

Secondly, sometimes there is a discrepancy between the analysis of national assessment results and recommendations that are promoted by assessment reports. For instance, the 2016 Grade 4 *Évaluation Nationale des Acquis* emphasizes teacher training needs in its conclusions, whereas this has not been supported by the data analysis conducted in this report. In addition, in some instances, recommendations do not differ greatly from one assessment cycle to another, or even from one assessment to another. For example, the NEA 2013 report borrows from international research and EGRA/EGMA[17] 2013 recommendations instead of putting more

17 This could be partly explained by the fact that RTI International produced both assessment reports. It should also be noted that the NEA test includes EGRA-like items. However, some conclusions and recommendations of NEA reports are taken directly from EGRA/EGMA findings.

emphasis on its unique findings and analysis, although the two assessments do not assess the same grades and the same subjects. Thus, although learning quality is not likely to change dramatically from one assessment cycle to another, it is critical that recommendations are consistent with the analysis conducted.

SEACMEQ reports provide good illustrations of precise, concrete, and well-targeted recommendations. They provide clear policy suggestions for national and regional-level officials, and assign a lead unit as well as a timeframe for implementing those suggestions. This ensures clear guidance and increased accountability of actors involved. An example of a policy recommendation can be found in the Namibia SEACMEQ IV 2014 report:

> Policy Suggestion 7: Given the status quo of teacher guides for mathematics and reading, regional education offices need to ensure that all teachers without guides obtain a copy to boost knowledge transfer. NIED should coordinate the distribution of teacher guides to the regions. (Namibia, 2015: 115)

Thirdly, as argued above, choices made when disaggregating learning data are very important. In addition to revealing potential disparities, they inform subnational education actors about the state of learning in their jurisdictions.[18] On the other hand, the level of data aggregation matters as well. As census-based assessment data provide information about all schools, they need to be aggregated to guide national-level planning and policy-making. However, according to one DNEA representative in Namibia, while the NSAT reports are prepared for each school, national and regional reports are

18 However, this depends on the assessment sampling design and whether it is representative at different administrative levels.

rarely produced,[19] which limits the use of data for central-level officials.

Because of a combination of these shortcomings, in Guinea, Namibia, and Senegal, multiple interviewees reported difficulties in extracting the relevant information from national learning assessment reports and identifying policy implications in their respective fields of intervention. For instance, according to a representative of the General Inspection of Education (Inspection Générale de l'Éducation [IGE]) of Guinea, despite a general understanding of the importance of learning assessment data, the department has difficulties exploiting them in their interventions, as there is no guidance and support to help them make the best use of assessment data.

3.1.3 The format of reports matters

> If a tree falls down in the forest but nobody is around to hear it, does it make a sound? … A fundamental question we need to stop and ask ourselves is: does a learning assessment matter if nobody understands or even reads the report? (Schwantner and de Chaisemartin, 2017).

As discussed above, the content of the report largely determines how final users understand and act upon findings. Nevertheless, our evidence also reveals that the format of the report is just as important for policy messages to be effectively conveyed, yet it is often not specifically designed for a general audience.

19 The research team did not have access to these reports.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

When developing assessment reports, it is important that authors consider the limited time and capacities that readers might have to read through full reports. This is even more relevant in countries where significant capacity issues exist. For dissemination to be effective, good practices indicate that reporting methods need to take into account the information needs of the target audience, their expected technical knowledge, and the most effective communication method (ACER-GEM; UIS, 2017). However, in all countries in the sample, respondents expressed dissatisfaction with the current design of assessment reports, as they are often too technical and not user-friendly, which limits the use of their data. A UNICEF Namibia representative explained:

> *The examiners' report is typically long-winded with a lot of narrative and text. This limits the uptake of information among school-level stakeholders. More innovative ways to think about how this information could be packaged and formatted to really speak to the end users will increase the utility and use of data. To do that, you need to ask the person. You must not make assumptions about the needs of the users, you have to engage with the users directly.*

While international assessment reports (in particular, PISA-D) provide ample information and meet all requirements in terms of the robustness and comprehensiveness of the analysis, they are rarely used, due to their high level of technical detail. Moreover, as mentioned earlier, decentralized-level officials also rarely rely on them due to a lack of data disaggregation at their level. In Senegal, a similar problem exists with the PASEC assessment; one of its representatives stated that they had taken steps to draft their reports in less technical terms to facilitate their use by the Ministry of National Education (Ministère de l'Éducation Nationale [MEN]).

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

Indeed, as suggested by DFID (2011), other tailor-made dissemination means should be used to communicate assessment findings in a non-technical and easy-to-understand language. However, limited capacities and funding do not allow a set of alternative dissemination products to be developed. The difficult task of the sole report produced is therefore to deal with the irreconcilable needs of a variety of actors with different capacity levels, expectations, and responsibilities. Consequently, diverging opinions about the quality of the reports often exist in countries: a statistically proficient audience would judge them not thorough and robust enough, while most stakeholders would find them barely readable. A couple of interviews illustrate this:

> *People don't even read the content, and when you read it, if you're not a statistician, you can't know what it means or what to do.* (Development Strategic Office [BSD] representative in Guinea)

> *Statistically minded people were saying the analysis is limited because the average does not say anything about the distribution, and they were arguing that the standard deviation and other elements need to be explained. However, others were saying that if the analysis is too statistically difficult, people will not be able to understand and might distrust it.* (UNICEF representative in The Gambia)

It is important to find the right balance between depth of analysis and reader friendliness, so as to maximize the assessment findings' uptake. The SNERS 2016 report provides hints about the way it could be achieved, as it starts with infographics summarizing the main findings. The manual of good practice of learning assessment indeed argues that 'reiterating the core messages increases the likelihood that audiences will engage with the reporting products'

(ACER-GEM; UIS, 2017: 44). Moreover, more technical details can be included in a dedicated appendix addressing an expert audience. According to ACER-GEM; UIS (2017), the assessment agency may also offer analytical services to the public whenever possible, to address points that are not covered in the final reports and ensure 'the data is widely used, and not dependent on stakeholders' high level of technical expertise' (ACER-GEM; UIS, 2017: 48).

> it is important that authors consider the limited time and capacities that readers might have to read through full reports.

3.2 Report dissemination an essential follow-up to assessment results

The implementation of assessments alone will not bring desirable system-wide changes in terms of learning outcomes if their findings are not widely shared and used. Indeed, '[w]here findings have been widely disseminated, they have raised public awareness, making education an issue on the public agenda' (DFID, 2011: 15). In this section, we look at how assessment findings reach key education stakeholders at national and subnational levels, and what factors hinder or improve dissemination processes.

Our research suggests that dissemination is an often-neglected phase of learning data management,[20] and it is also heavily affected by budgetary issues. In many countries of our study, dissemination is not part of a coherent strategy

20 The management of learning data refers to the design of an assessment framework, development of data collection tools, data collection, data management and archiving, data analysis, dissemination, and use of learning data.

that includes follow-up activities on the use of assessment data, and, in practice, multiple inconsistencies emerge in the way data are disseminated, as further examined below (e.g. certain actors are not reached, data release is delayed, and there is a lack of financial resources for implementing planned activities).

3.2.1 A consistent dissemination strategy can facilitate sharing of assessment results

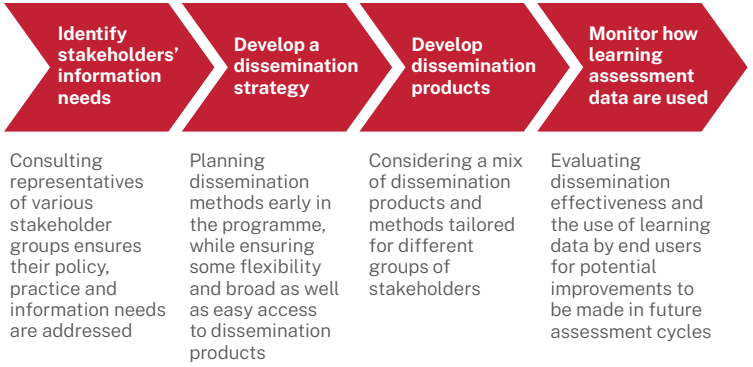
A Teaching and Learning: Educators' Network for Transformation (TALENT) policy paper on the effective use of large-scale learning assessment data explains that '[t]he purpose of the dissemination strategy is to forecast the availability of general and specific results to actors, stakeholders and the public, between the completion of the analyses of results and the next round of the assessment' (TALENT, n.d.: 6; Robertson, 2018). The literature suggests that it is important to develop a strategy identifying dissemination channels, units in charge of dissemination, and end users, not only because it maximizes stakeholder engagement, but also because these factors will affect decisions on subsequent analyses and budgetary provision (ACER-GEM; UIS, 2017; DFID, 2011) (see *Figure 5* for more information on key elements to consider for a comprehensive dissemination approach). An assessment strategy could take into account the following points (Robertson, 2018):

- identification of stakeholder information needs;
- resources available and persons responsible;
- evidence about what dissemination methods are most effective;
- possible policy implications of assessment results;

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

— anticipation of likely issues or questions;
— main dissemination product and secondary products;
— availability of technical information and dates of release;
— evaluation of impact (i.e. monitoring how different dissemination products and assessment data are used over time).

Figure 5. Key elements to consider for a comprehensive dissemination approach



Source: ACER-GEM; UIS, 2017.

In most countries, officially defined dissemination strategies for sharing assessment results do not exist. However, Zambia and Namibia have examples of more structured dissemination processes. The NLAF defines the recommended information flows between actors at all levels and later assigns the responsibility of NAS dissemination to the ECZ; nevertheless, it does not specify dissemination procedures to follow. In Namibia, circulars specify means of dissemination for each NSAT cycle. For NSAT 2017, a circular[21] was sent together

21 Circular DNEA 25-2018 (Namibia, 2018) 'Dissemination of Grade 5 Baseline Standardized Achievement Tests Reports'.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

with the NSAT reports to subnational offices reiterating the main purpose of these reports and providing details about the dissemination process. It designates the regional examination officials (REXOs) as facilitators and coordinators of the NSAT delivery in their respective circuits and schools. It then urges school principals, heads of departments, and subject heads to further share and analyse results with their staff in order to design appropriate remedial activities.

Mechanisms that ensure users have received, read, and understood the reports are essential for their ability to act upon findings, and yet they are often missing. Even in countries where dissemination seems to be effective overall, feedback mechanisms to schools, districts, or regions on suggested corrective actions and their effective implementation are often missing or ineffective. Effective feedback loops could make it possible to determine whether the use of the data is taking place as intended and, if not, what the reasons are for this. Changes can then be made to adjust dissemination strategies to ensure better targeting of policy and information needs, and to provide adequate support on the use of assessment data. A representative from the Gambian Project Coordination Unit emphasized the importance of the follow-up on assessment results:

> *Just publishing data is not enough; it is important to do some deep research and ask questions like: are the schools really using it? Are schools using the recommendations of the analysis? What has been done with the data? We need to see whether data are really helping us to move forward. It is not only about publishing it, but you need to do a follow-up on the recommendations and whether they are implemented.*

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

The importance of a dissemination strategy has been recognized in some project countries. For instance, in Senegal, the development of a dissemination strategy will be part of a performance contract established between the INEADE and the General Secretary (*Secrétariat Général*). In Zambia, ECZ and DPI have started working together on a dissemination plan to coordinate and expand dissemination efforts and implement the NLAF more effectively. Moreover, a new strategy has been put in place to engage a wide range of stakeholders in the discussion of assessment results and implications for policy.

3.2.2 Inconsistencies in data dissemination

In most countries, dissemination is often ad hoc and not targeted. Dissemination activities are not yet organized in a systematic way, which sometimes negatively affects the use of assessment findings.

Although learning data are often easily accessible for actors at central level, the chain of dissemination seems to be disrupted at decentralized level. When preparing ESAs and ESPs, central level officials often receive assessment reports from their colleagues in charge of assessments. In Namibia and Senegal, many respondents reported that actors at central level (e.g. curriculum unit, directorate of primary education, planning unit) are informed about the latest assessment results. In Ghana and Guinea, learning data are also considered to be relatively easily accessible at central level, though not systematically.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

In contrast, more challenges arise at decentralized level. Assessment reports sometimes reach the first subnational level (i.e. provinces/regions), but more rarely the lower administrative levels, including schools. This situation excludes a significant number of end users who are in a position to take concrete action in education systems and impedes the effective use of data.[22] Interviewees at subnational levels had contrasting views about results being shared at their level, which reflects the non-systematic nature of dissemination. In Ghana, Namibia, and Zambia, decentralized-level officers noted that provinces, districts, and schools do not receive analysed assessment results at every assessment cycle. In Senegal, dissemination is not as targeted and broad as at central level. In fact, several Senegalese actors said that school inspection offices at the provincial level (Inspection d'Académie) must have assessment results, but they do not systematically relay them to school inspection offices at district level (Inspection de l'Éducation et de la Formation [IEF]) and schools. In Guinea, dissemination at decentralized level remains very limited. In addition, in Zambia, the ECZ is aware of dissemination issues as it observed that '[s]ubsequent large-scale assessments conducted after 2012 have mainly adopted a national- and regional-based dissemination model, which largely leaves out the key stakeholders in the use of assessment results' (Zambia, 2019a: 14). As mentioned by a decentralized-level official in Zambia, the issue has concrete implications in the classrooms:

22 As noted earlier, it is important to consider that, depending on the assessment sample and data disaggregation modalities, the assessment analysis might not provide equally relevant information to actors at all administrative levels.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

> *There is a need for a proper system to have the reports reach the teacher in the schools. … For example, in Grade Five, year in year out there is a problem with fractions, and this problem was unknown to teachers and yet they are the ones who teach. For me, feedback should be instant.*

Similar dissemination issues emerge with cross-national assessments. In Namibia, although SEACMEQ aims for wide dissemination, most regional officers indicated that they do not have access to SEACMEQ reports or have not received any data for a long time. For example, SEACMEQ III results were not disseminated in the country.

Dissemination of assessment data is strongly affected by a lack of financial resources, as it is often the first area to experience financial cuts; this has a direct negative influence on the accessibility of learning data and, in turn, their use. In Namibia, Senegal, and Zambia, it was reported that budgeting issues have slowed down the intended implementation of dissemination activities. For instance, in Senegal, development partners supported the dissemination of brochures presenting PASEC 2014 findings, but only four regions could be covered. In addition, a lack of funds affected data-sharing from regions to districts and from districts to schools. In Zambia, dissemination activities are facing budgeting challenges, especially for the ECZ. In Namibia, budgetary constraints strongly limited contact with schools, especially remote schools, and created delays in delivery of results to students. Workshops and meetings to disseminate NSAT and SEACMEQ results were restricted and this adversely impacted information sharing. The literature reports a similar issue with this assessment in Uganda, where limited resources affect the timely processing of results and an inadequate dissemination budget 'only allows it to work

with senior figures in government and NGOs' (Elks, 2016: 27), leaving teachers unaware of assessment results.

Assessment data arrive too late to inform planning. In all countries, timeliness was identified as a significant challenge – though to a different extent – for the use of learning assessment data in planning, especially at decentralized level. Timeliness issues were voiced strongly in Namibia, Senegal, and Zambia. In The Gambia, delays in the data analysis chain slow down the development of SIPs. The dissemination of SEACMEQ IV results was delayed considerably in both Zambia and Namibia, which made data obsolete when they were finally released. In Namibia, the data generally arrive months late for planning at regional level. One chief education officer said that '[w]e can only use the data from the previous year, not the data for the year for which we are doing the planning', and this issue is also evidenced in Uganda: 'S[E]ACMEQ requires greater resources to process the data quickly if the benefits of taking part in the assessment are to be realized. If the decision is taken to participate in an exercise of this cost, then sufficient resources must also be allocated to the final stages of the process, to ensure value for money' (Elks, 2016: 28).

Flexibility in data dissemination is needed. As argued in the Principles of Good Practices in Learning Assessment, 'flexibility is also needed as new stakeholders are identified, as resources become available, as evidence is gathered about what dissemination methods are most effective, and as other possible policy implications of the assessment results become apparent' (ACER-GEM; UIS, 2017: 44). In some cases, flexibility in choosing dissemination modalities adapted to the context may alleviate budgetary repercussions on dissemination and better respond to stakeholders' practices and local contexts.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

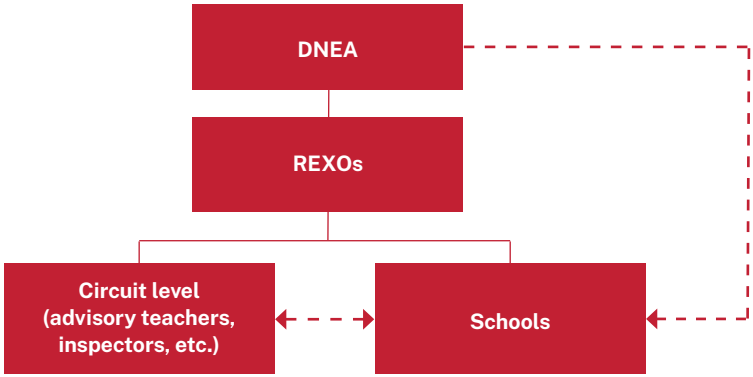
In Namibia, actors seem to disseminate assessment reports through the means they consider the most practical, even if they are not specified in circulars on data dissemination. This allows data dissemination to be more effective. The solid lines in *Figure 6* show the dissemination path as defined by circulars, while the dashed lines show other ways through which assessment results reach schools. Regional officials reported different methods used to share the results: social networks, emails, hard copies, or in person. For instance, a DNEA representative explained that they could send the results directly to schools by email when they have access to internet facilities. However, some arrangements may introduce asymmetric access to learning data. In Ghana and Guinea, while informal means of dissemination are sometimes the solution to the non-systematic sharing of assessment reports, it might penalize actors with less developed networks within MoEs. In Guinea, one respondent explained that without a network within the ministry to request information, it would not be possible to access the data.

3.2.3 A mixed-method dissemination can facilitate uptake of assessment results

Beyond providing information, dissemination methods should also support actors' capacity to understand and make effective use of assessment results. Moreover, the literature underlines the necessity of developing a variety of dissemination products (DFID, 2011; TALENT, 2020). However, two dissemination methods are commonly used in our sample countries: assessment findings are shared either in meetings with education stakeholders and/or in hard/soft copies of assessment reports.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

Figure 6. Dissemination channels of the NSAT results in Namibia



Source: Developed by authors.

Dissemination meetings offer the advantage of ensuring participants are actively exposed to the information. They can help trigger a discussion around the results and raise awareness around learning issues. In Namibia, the DNEA used to organize meetings throughout the country for the release of the NSAT results. Dissemination meetings are also held at national and provincial levels in Zambia. However, the effectiveness of these meetings depends on the quality of exchanges, which varies. For instance, an interviewee from a regional office in Namibia explained that interactions during meetings with central-level officials on learning assessments were 'merely informative' and did not allow for a critical debate. In addition, as related costs of these meetings (e.g. transport) can be high, budget issues often limit attendance.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

The distribution of hard/soft assessment reports is another important dissemination modality. Due to budget cuts, dissemination in Namibia now relies heavily on the distribution of the NSAT reports. Likewise, in Guinea, since there are no meetings where assessment results could be shared and discussed among education stakeholders from different jurisdictions, dissemination relies on the distribution of assessment reports.

The study also illustrated interesting examples with a variety of means of dissemination that strengthened the uptake of learning data. For instance, in Ghana, multiple dissemination activities have been organized to share NEA, TIMSS, and EGRA/EGMA results (e.g. presentations for key stakeholders, seminars, training workshops, and district cluster forums). For EGRA and EGMA 2013, a forum brought together district-level education stakeholders and established district advocacy teams that would further disseminate results and organize actions to support early grade learning for each district (EGRA/EGMA report, 2016). Seminars were also held at the MoE to share TIMSS results.

According to the Principles of Good Practices in Learning Assessment, 'a mix of dissemination methods and products probably best addresses the information needs of different stakeholders' (ACER-GEM; UIS, 2017: 44). Acknowledging the diversity of education stakeholders that assessment results intend to reach, an effective dissemination pattern ideally relies on a variety of methods that should go beyond a combination of meetings and assessment reports, although doing so ultimately depends on available resources and capacities. Zambia offers interesting insights into how different methods can be used to disseminate assessment

findings (see *Box 3*). Additionally, the release of various dissemination products can be gradual to maintain interest and momentum in the assessment (ACER-GEM; UIS, 2017).

BOX 3

Dissemination of assessment data in Zambia (before 2012)

Zambia developed a variety of dissemination materials for the National Assessment Survey findings. Two reports (the main report and a summary report) with narrative, graphic, and statistical information were shared with the ministry staff (READ, 2010). Leaflets with action points and recommendations, as well as posters and policy briefs were then addressed to school, community, and education administrator levels (READ, 2010; ECZ, 2019). With the support of the World Bank 'Russian Education Aid for Development' (READ) project, Zambia produced an educational film on the results of the 2012 NAS for grade 5 to be used in pre-service teacher education (Zambia, 2019a). Dissemination meetings were then conducted at provincial and district level throughout the country, while the general public learnt about the main findings through the radio, television, and media releases (UIS, 2015).

Before 2012, major factors for the successful dissemination of NAS included attention on the accessibility and appeal of the materials. According to the Zambia Self-Diagnosis (2010), these aspects were considered after representatives of target users were invited to provide feedback during a workshop. Perhaps just as importantly, '[s]upport is offered in the interpretation of results through INSET [*In-Service Education and Training*] providers' (READ, 2010: 65).

3.3 Learning assessment data as a component of an information ecosystem

In this section, we look at how learning assessment data fit into the EMIS and broader information ecosystems. Our evidence suggests that, very often, learning assessment data are not part of EMIS, although in countries where it is being reformed (see below), efforts are made to incorporate these data. The unavailability of learning data in EMIS reduces their accessibility and impedes their more effective use. Having learning assessment data in EMIS is even more important in countries where information flows between units are weak and other means of communication are absent.

3.3.1 Integrating assessment data in EMIS: why it is important

The development of a solid EMIS,[23] together with a robust learning assessment system, is 'core to the country's capacity to produce and monitor education data and evaluate whether children are learning' (GPE, 2019: 2). Development partners played a key role in supporting the development and strengthening of EMIS in low-income countries. Ninety-five per cent of GPE grants active in 2019 had support to learning assessment systems or EMIS as one of their components.

It is important to include learning assessment data in EMIS to facilitate the availability of relevant and timely data for planning (UNESCO, 2003). Moreover, it can help relevant units in the MoEs, researchers, as well as development partners

23 An education management information system (EMIS) can be defined as 'a system for the collection, integration, processing, maintenance and dissemination of data and information to support decision-making, policy analysis and formulation, planning, monitoring and management at all levels of an education system' (GEM Report Team, 2008: 101).

who wish to conduct secondary analyses on assessment results, to have easier access to these data (UIS, 2018b).

3.3.2 Learning assessment data often missing in EMIS

Project countries have embarked on reforming EMIS, and, in most, the integration of learning assessment data is ongoing.

Namibia has committed to better management of the education data with the support provided by UNICEF. At the time of data collection, the restructuring of EMIS and the creation of a related policy were underway. A UNICEF representative explained that even though the EMIS report 'reproduces all the DNEA outputs faithfully', the M&E unit, responsible for EMIS, did not deal with the raw data from the national assessment (i.e. NSAT) and the EMIS does not contain learning assessment data.

In ***Senegal***, the EMIS contains continuous assessment and examination results, as well as subnational standardized assessments (ESH) grouped by district levels. In addition, other administrative and pedagogical management applications (Planète H, Planète Examen, StatEduc2) allow easy access to a wide range of data that are particularly useful for headteachers and planners. Nevertheless, none of these databases contains SNERS (i.e. national LSA), EGRA, or PISA-D datasets.

In ***Zambia*** and ***Ghana***, learning assessment data are currently scattered in different databases. In Zambia, the ECZ database for assessment results is not part of the EMIS, whereas some data are also stored with third parties (i.e. DevTech and MSI for EGRA/EGMA data), other units (i.e. DPI for SEACMEQ data),

or the Zambian National Data Centre. Moreover, some of the more recent NAS and EGRA/EGMA data are available online, but this is not systematic throughout assessment cycles. Some interviewees expressed the need for a central data pool. As part of the Zambian Education Enhancement Project (ZEEP), the MoE has engaged in a World Bank-led project to enhance the EMIS and harmonize data. Nevertheless, the terms of reference of the project do not refer explicitly to incorporating LSA data (Zambia, 2019b). The situation in Ghana is very similar. Although the most recent assessment reports have been available online, assessment data are scattered in databases of different agencies and divisions of the MoE. Moreover, the EMIS does not contain NEA nor EGRA/EGMA data. However, the Ghana Accountability for Learning Outcomes Project aims to create an integrated dashboard that includes learning assessment data.

> “ Dissemination meetings offer the advantage of ensuring participants are actively exposed to the information.

In contrast, in The Gambia considerable improvements in the EMIS development included the integration of learning assessment data; the EMIS has been collecting the NAT data since 2011 (Senghor, 2014). In addition, the Assessment Policy (2015–2022) has identified problems with the strategic storage of data and delegated this responsibility to the planning unit (i.e. PPARBD) for storage, analysis, reporting, policy advice, research, publication, and dissemination. Irving and Mitra (2019) underline the successful implementation of a programme called ‘Results for Educational Achievement and Development’, co-financed by the GPE, which aimed to improve the production and use of data:

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

> The programme linked National Assessment Test (NAT) results across different years to allow comparability of results over time and linked these data to EMIS data, education sector human resources data, and data from regional education departments. Country-level stakeholders were able to use this comprehensive set of data for regular, evidence-based sector reviews and implemented recommendations from these reviews to improve sector plan implementation. (Irving and Mitra. 2019)

In several countries (Ghana, The Gambia, Senegal, and Zambia), our research reveals that key departments consider EMIS as the primary source of information for the development of strategic plans as well as routine planning activities. The absence of learning assessment data in EMIS might therefore have a negative impact on the use of learning assessment data for these purposes and create a gap in the use of learning assessment data as compared to other evidence. Nevertheless, this does not necessarily mean that EMIS databases should match students, schools, and teachers with the corresponding test results since, if made public, this information might increase the stakes of assessments and foster competition between schools and teachers. For many purposes, it may be sufficient to match test scores through more aggregated criteria. It depends on a political decision regarding learning assessments and the way their data are intended to be used.

In addition, using each piece of education data in isolation could be misleading if the data are not cross-checked with other evidence. Evaluating learning quality through a single learning indicator may have repercussions, such as neglecting the social, economic, and contextual dimensions of education quality (UNESCO, 2019). Hence the importance of looking across various data – which EMIS facilitates.

3.4 Capacity-building as an internal engine driving use of assessment data

When learning assessments are introduced in countries with low national technical capacities[24] and without a supportive institutional framework, national education systems encounter challenges in producing and using data. Some assessment-related activities exceed existing capacities. Capacity development is therefore key to sustainably improving the use of learning data and their ownership.

3.4.1 Capacities for producing and using learning data remain low

Conducting large-scale learning assessments and ensuring effective use of their data are resource-intensive tasks that demand strong financial, human, and technical capacity. It involves a large set of highly technical and complex tasks, such as designing and constructing assessment instruments and background questionnaires, using sampling techniques, organizing testing in schools, data preparation and analysis, and, finally, report writing as well as dissemination (DFID, 2011). The use of learning data also requires statistical understanding from actors that are not necessarily comfortable with quantitative analysis. These processes are therefore capacity-intensive: 'Large-scale assessments involve an enormous amount of organization of processes, staff and

24 The study examined the capacities of officials in charge of the production and use of learning data based on what they reported in the interviews. It is therefore important to introduce a certain degree of caution, as the study only allowed *perceptions* of actors' capacities to be gathered, and these perceptions might contain some bias. In addition, the study did not look into school-level capacities for using learning data in depth, although multiple interviewees reported them as being low in all countries.

logistics. There are also often very demanding timeframes involved, with many finish-to-start task dependencies' (ACER-GEM; UIS, 2017: 12).

It is important to differentiate between capacities to produce and use learning data that require different sets of skills. Whereas assessment teams need strong *statistical* skills and knowledge in sampling procedures, psychometric quality, validity, and reliability of assessment instruments and the application of appropriate analytic methods (GPE and ACER, 2019), decision-makers need to focus on the interlinkages between assessment data and policies, and therefore require skills for understanding data. Although assessment reports sometimes provide ready-made recommendations based on assessment results, it is important for planners and policy-makers to have capacities to understand the basis for the analysis of assessment results. The study revealed that a lack of capacities for using learning data is more often overlooked in project countries.

This issue is exacerbated by a lack of guidance for actors on how to make the best use of the information assessments provided in their everyday activities, and a general lack of statistical skills within MoEs. In Guinea, several interviewees felt that statistical skills were missing among data users. Assessment reports failed to trigger a critical discussion on student learning and influence political agendas, partly due to a lack of users' capacities. In addition, one respondent emphasized that they were not well informed of how assessment results can be used in their work, and that guidance on this is very much needed. In Senegal, interviewees also mentioned the limited capacities of data users.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

Although there have been many capacity development activities conducted in the project countries, study results and the existing body of literature (see, for example, Wagner et al. 2018) confirm that overall capacities for producing and using learning data remain low. However, capacity needs vary by country and by administrative level. Some countries managed to develop solid assessment teams with relatively strong capacities that only need occasional support in specific and more complex assessment tasks, such as the INEADE in Senegal and the ECZ in Zambia. In others, capacities need to be substantially improved (e.g. Guinea). In addition, technical skills were weaker at decentralized level in all project countries. A number of recurring capacity development needs have emerged in the study:

- advanced statistical analysis, especially in terms of econometric techniques for estimating the influence of different background variables on student learning, data comparability over time, as well as statistical sampling and survey methodology;
- data interpretation and translation of findings into policy recommendations;
- use of statistical software for data analysis (e.g. STATA, R, SPSS);
- data and database management;
- dissemination and communication strategies;
- report writing, ensuring its accessibility and reader friendliness.

In project countries, the lack of skills in statistical analysis is a more general issue affecting the entire education sector. This creates situations where a small group of qualified individuals has to assume a considerable share of responsibilities. In

project countries, statistical skills are distributed unevenly across and inside units. A MoBSE official in The Gambia highlighted this issue, adding that it is key to improve the sharing of existing knowledge and skills: 'You know the 20-80 rule, where 20 per cent of employees do 80 per cent of the work. That 20 per cent should share their expertise better.' This issue was particularly salient in Ghana and Guinea, and to a certain extent in other project countries, as a small number of qualified people in assessment teams have to take on considerable responsibilities in large-scale national and regional assessments, which leaves them depleted. A review by the Organization for Economic Cooperation and Development (OECD) also identified this issue in Zambia:

> The main bottleneck is with the specialized technical staff, who are few in number and tend to be responsible for a wide variety of tasks, including management, data processing, analysis, document preparation and reporting. These resources are not efficiently used, because a large portion of their time is spent on routine clerical tasks (e.g. communication and document preparation). Peer-to-peer meetings are typically sacrificed in favour of preparation for meetings with senior managers or administrators, so there is also inadequate opportunity for staff to collaborate effectively. (OECD, 2015: 22)

A lack of national capacities in some cases leads to the outsourcing of more complex assessment activities, which interrupts the national cycle of data management, reduces national ownership of data, and prevents countries from developing assessment capacities further. In Guinea, the INRAP, which is responsible for EGRA, often outsources the data processing, analysis, and writing of the final report. In Ghana, the NEAU also outsourced data analysis and report writing.

Assessment teams also often suffer from and are particularly sensitive to high turnover rates. A very limited pool of candidates with the necessary skills in countries further exacerbates this issue. A high turnover of assessment teams leads to a lower return on investment in capacity development, as trained people leave the system. In Zambia, the ECZ faced difficulties replacing and upskilling its human resources after more than 20 years of conducting LSAs. In Namibia and Senegal, the assessment team turnover rate is high. In Namibia, this has created a constant need for capacity development since 'people come, they get trained, and they go', as noted by a DNEA representative. In Senegal, many people trained by the INEADE end up leaving the service, including after training. Consequently, the number of specialists working in this service is currently not sufficient to take on the workload required. The study cannot provide a comprehensive explanation for this trend, but one reason mentioned by interviewees suggests that, as there is a limited number of people with high-level statistical skills in these countries, they are in high demand across sectors.

3.4.2 To improve the use of learning data, capacity development should be the focus

Focusing on developing national capacities can improve the use of data and data ownership in a sustainable way, as was the case in The Gambia. The national capacity development became an internal catalyst not only for increased ownership of the assessment system, but also more effective use of data. The permanent secretary at the time of the NAT development insisted on strengthening national capacities instead of heavily relying on external consultants. The AU was created in 2015 and has expanded gradually since then. Team

members participated in several training events and the World Bank consultant assisted the team with more complex tasks, progressively transferring skills and knowledge. In The Gambia, the AU is therefore becoming a technically sound focal point for assessments and brings together other departments involved in assessment-related tasks, which facilitates the production and use of learning data.

Regional and international learning assessments contribute to national capacity development through dedicated activities. SEACMEQ, PASEC, and PISA-D include capacity development as one of their objectives and dedicate resources accordingly. For instance, a capacity needs assessment (CNA) is a prerequisite for countries to participate in PISA-D. Senegal and Zambia had to design capacity-building plans based on CNAs and the OECD supported their implementation.

The positive influence of participation in international or regional assessments on national capacity development is also supported by the available literature (Lockheed, 2013; Fischman et al., 2019). SEACMEQ studies focused on capacity-building in the area of educational policy research and had a programme of intensive training workshops that introduced local researchers and planners to all phases of the study (Paviot and Saito, 2015). Every participating ministry reported positively on the benefits of capacity-building; this impact was reported to be the highest in Botswana, Lesotho, the Seychelles, and Malawi (Murimba, 2005). PASEC countries also improved capacities in test construction and in the design and execution of surveys (Kellaghan and Greaney, 2003). Some studies also highlight their influence on the overall development and/or improvement of national assessment systems (Breakspear, 2014; Murimba, 2005).

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

Capacity development activities are key to ensuring the sustainable development of assessment systems and effective use of their data. Yet when it comes to participation in cross-national assessments, procedures and tools are often prioritized to bring about expected changes that, in reality, cannot take place without capacity development. The latter is also key to developing self-reliant teams ready to innovate and adapt assessments to their contexts, ensuring their impartiality from political influence:

> It is worth noting that human capacity appears underemphasized in the current literature on education data. In particular, human capacity to bring about innovation within individual countries seems underemphasized. Instead, much of the focus falls on tools in the form of manuals and standards. These tools are important but do not guarantee, on their own, that the necessary human capacity will be built. Cross-national assessment programmes have created networks that have facilitated country-specific capacity-building. Yet the processes within these programmes are largely premised on a model where innovation and advanced technical work, for instance with respect to sampling and psychometrics, occurs in one place, while each country follows a set of instructions. The problem with insufficient innovation (as opposed to imitation) in individual countries is that country-focused use of the data which emerges from the cross-national programme is often limited, as is capacity to design national programmes. Moreover, weak technical capacity in a country might mean that national assessment systems are influenced by political interference, which is a real risk in an area such as assessments. (UIS, 2018b: 23)

“ Capacity development activities are key to ensuring the sustainable development of assessment systems and effective use of their data.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

3.5 Smarter investments needed to ensure cost-effectiveness and improve use

3.5.1 Cost-effectiveness of learning assessment systems is key to actors' buy-in and support

Although some authors argue that learning assessments are inexpensive if compared to their potential positive effects on the education system,[25] their absolute costs can be significant, especially for low-income countries. External partners entirely or largely cover these costs in the project countries due to a lack of national funding, which also reflects a more general trend:

> A general trend of externally financed learning assessments is observed for both national and cross-national assessments, and particularly, in developing countries. For example, at least two-thirds of national learning assessments are funded with the support (in part or fully) of foreign or external donors (UIS, 2019a), such as the GPE, UNICEF, USAID and the World Bank. For PASEC 2014 and beyond, the CONFEMEN funds half of the cross-national assessment while the government will need to finance the rest to participate. (UNESCO, 2019: 52)

The UIS has been advocating for improved assessments and better use of their data, considering their potential benefits: 'Learning assessments are among the least expensive education reforms, typically costing far less than building schools or hiring teachers. When correctly implemented, learning assessments can be used to monitor learning for all and, most importantly, to improve learning. Assessments can be among

25 Different benefits can be considered, such as diverse economic, political, technical, and socio-economic rationales for participation and non-participation and 'debates about whether ILSA participation is "worth it" must take account of the diverse purposes of participants in these assessments' (Addey, 2019: 1).

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

the most cost-effective reforms a country can implement' (Ramírez, 2018: 5). Nevertheless, absolute costs can be significant for developing countries, especially considering conflicting priorities in education systems. Although the Foreign, Commonwealth & Development Office of the United Kingdom (FCDO, previously DFID) also recognized high opportunity costs in not undertaking assessments, as they are key to guiding future investments, it also emphasized high overall and opportunity costs when undertaking assessments, which makes it key to improving return on investment:

> The absolute costs may be significant, especially in countries where the non-salary budget is small. The cost of a national assessment will be sufficient to buy many textbooks or employ more teachers, making it unattractive to policy-makers. The value of national assessment increases as the time series of comparable data builds up. Therefore, the decision to allocate funds should be considered as an ongoing commitment over the medium and long term, not a one-off expenditure. (DFID, 2011: 4–5)

Because assessments are expensive endeavours, it is so much more important to invest smarter and strategically in order to improve the use of data. If learning data are not subsequently used and do not bring the expected benefits of informing education systems, the cost-effectiveness of assessment systems decreases substantially, as does support from national actors. Some of those interviewed expressed their concern about the 'we already know' trap, and the opportunity cost of conducting assessments over implementing other policies. In some instances, assessment results repeatedly demonstrate low learning outcomes and stagnating results over the years, without necessarily providing new insights on the quality of learning. As one Senegalese official put it: 'We have to use these results to guide programmes. But we have the impression that we're doing the same programmes

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

whatever the learning results.' A Guinean official expressed a similar concern: 'Assessments are too expensive for their usefulness … in terms of usefulness we can say that it's expensive because, given the use made of it in relation to the resources committed to it, we would have preferred to allocate it to something else.' Similar doubts about the opportunity costs of assessments can also be found in the literature:

> Regular monitoring of learning achievement is important, but I just don't understand the rationale of development partners' push for implementing/piloting international/regional assessment instruments. Despite knowing the fact that, for years, the education system continues to struggle to implement the instructional hours stipulated in the curriculum, teaching and learning are being delivered by under-qualified and under-supported teachers in crowded classroom settings, with shortages of textbooks and learning materials, this is going forward. Further, the results from existing learning assessments, examinations, etc. are already pointing to the fact that massive efforts and investments are needed in improving reading and mathematics from early grades onwards. So why get distracted or distract? I would rather use the scarce money to provide … teaching guides to the teachers. (UNESCO representative, interview conducted March 2017) (Auld, Rappleye, and Morris, 2019: 10)

The issue of cost-effectiveness of assessments and their opportunity cost is even more problematic at secondary level for low-income countries: 'Surveys in India show, for example, that 50 per cent of fifth graders in India cannot read at Grade 2 level (Pratham, 2015). Unless that is remedied, there is not much that can be expected from 15-year-olds. Engaging in a costly and sophisticated test of 15-year-olds, a large proportion of whom are not even in school, is wasteful, and the scarce resources could be put to better use' (Addey and Gorur, 2020: 14–15). Although the study did not allow this issue to be addressed in depth, similar concerns arise in

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

sub-Saharan Africa countries that also struggle to improve early grade literacy and numeracy. In Senegal and Zambia, limited use of PISA-D results for secondary level has been observed. Senegalese officials expressed concern, saying the 'assessment does not reflect the situation in the sector'. However, a more detailed analysis of the issue would be needed to confirm results and better understand the causes.

In the light of these arguments, a provocative question arises: do we need to continuously reinvest these amounts of money to prove the same lessons over and over again? As one of the senior officials from Uganda put it: 'What is the value of showing that children are not learning for the fourth year in a row?' (Elks, 2016: 20). The answer, as documented by our study, is to support more strategic investments in assessments and a broader vision of the use of data. This also includes 'optimizing' learning assessment systems to avoid unnecessary overlap and incoherencies, as analysed in *Chapter 2*. It is therefore important to ensure continuity of assessments, improved national capacities and communication, to ameliorate the use of data and to increase return on investment:

> Investment is more likely to take place if the benefits are clearly communicated. In other words, a stronger emphasis is needed on the demand for and utilization of data, not simply supplying data [Ramírez, 2018]. This requires thinking differently and more broadly about processes around data. For this, human capacity is needed, both with respect to broad strategic thinking around data and also with respect to very specific skills. There is also a need for better technical documentation to guide countries. The challenge is to find the most cost-efficient, fit-for-purpose way of producing learning statistics. (UIS, 2018b: 23)

3.5.2 Financial constraints hinder smooth implementation and negatively affect use of data

As national financial resources are often limited and external funding is not always stable, obstacles to smooth implementation of assessments emerge and, in some cases, may even lead to their cancellation. For example, in Namibia, the EGRA assessment has not been conducted since its initial implementation in 2012 due to a lack of financial and human resources. Although the process had started, budget cuts impeded the development of EGRA tools in all 14 curriculum languages. Moreover, the NSAT was not conducted in 2019 due to budgetary constraints. Budgetary issues have also caused delays in PAQEEB and SNERS implementation in Senegal in recent years. In Zambia, financing shortfalls led to the partial implementation of planned EGRA and EGMA 2016 rounds. Moreover, although ESPs and programme documents foresee funding for learning assessment activities, many interviewees drew attention to the inconsistency between the planned budget and what is actually used for learning assessment activities.

The study also revealed a frequent misalignment between the arrival of the budget and the implementation of assessments. A release of funds from central to decentralized level is often delayed, as was the case in Ghana. At decentralized level, multiple interviewees noted a lack of financial resources and delays in funds' arrival, which consequently led to delays in data collection. In The Gambia, an interviewee stated that, as the financial support comes from international donors, sometimes there are delays in accessing the funds.

One Senegalese official also mentioned a similar problem: 'Sometimes the budget is available but the assessments have not yet been planned and this creates a delay in implementation. Conversely, sometimes assessments are ready to be conducted but the budget is late in coming.'

Other issues related to infrastructural needs for assessment implementation and data analysis were identified in project countries:

- logistical issues in accessing schools, issues with transportation, i.e. vehicles, devices;
- lack of internet connection;
- lack of computers and other tools for data collection.

3.5.3 Difficulties in implementing assessment recommendations due to financial constraints

Policy recommendations put forward by the assessment results analysis cannot always be implemented due to a lack of financial resources. Although some policies promoted by assessment findings are well received by decision-makers and find their ways into educational strategic plans, our study demonstrated that their approval or implementation is often hindered by limited resources. Officials in Ghana illustratively emphasize that: 'even if we know what and how we need to fix, we might not have the resources to fix it and to implement specific strategies. As there is a lack of resources everywhere, this also impedes the use of learning data for finding solutions'.

3. Dynamics between barriers and conducive conditions for the use of learning assessment data

The agendas of international partners influence which assessment recommendations are implemented with the support of their financing. This link is even stronger when they are the ones financing assessments, making it more likely for them to follow up with a project based on assessment results. For example, the influence of EGRA results on policies is clearer because they are often followed by internationally funded reading projects (e.g. Ghana, Senegal, Guinea, and The Gambia). In addition, those interviewed in Guinea were open about the fact that international partners often influence policies with their funding: 'At the same time, there are political priorities (the political priority letter and the budget) that do not necessarily allow for a focus on the issues highlighted in the assessment reports. Policies are often guided by the priorities of the donors funding the projects.' Data on learning can suggest policies, but if there is no funding they cannot be implemented. A similar finding is confirmed by the available literature in other countries:

> The impact of EGRA on influencing policy is hard to isolate. A greater policy focus on early reading quickly followed from the first administration of EGRA in Uganda. This was largely because of the availability of funding for the SHRP programme from USAID, which provided teaching materials and training to schools. Identifying funding channels is clearly a challenge for the managers of other assessments. Government officials compare this experience to the challenge of finding money to respond to areas of improvement highlighted by NAPE. (Elks, 2016: 22)

It is therefore important to ensure recommendations arising from learning assessments are costed, financially realistic, and deliverable (Elks, 2016).

Chapter 4. The interplay of national actors: between miscommunication and misunderstanding

A literature review conducted at the inception of the project identified a lack of cooperation between national actors as one of the most common barriers to the effective use of assessment data (Raudonytė, 2019). This chapter examines the in-depth causes of these communication issues to help understand them.

4.1 Cooperation between national actors at central level remains ad hoc

4.1.1 Limited information flows and exchanges on learning assessment data

The systemic use of learning data depends heavily on the quality of information-sharing channels as well as collaboration between AUs and actors responsible for different policy areas. However, collaboration within MoEs is often limited and poorly institutionalized and does not facilitate a wider sharing and use of assessment data.

Our evidence shows that in most countries, collaboration and communication between units producing assessment data and their potential users, i.e. units in charge of curriculum, pedagogy, teacher training, and planning, are somewhat limited – regardless of institutional modalities for assessment teams (see *Chapter 2*). This in turn hinders the uptake of assessment data by actors with decision-making power in key policy areas. Collaboration issues could be observed in

4. The interplay of national actors: between miscommunication and misunderstanding

Guinea, Ghana, Namibia, and Senegal, whereas in The Gambia and Zambia collaboration was mostly reported as smooth. The study documented common practices of 'working in silos', observed, for example, in Ghana and Guinea, hindering collaboration. The study highlights some problematic or conducive aspects of collaboration, and their potential impact on the use of assessment data.

In most countries, collaboration and information flows between actors involved in the production and use of learning assessment data are not institutionalized. Few institutional mechanisms (i.e. committees, networks, peer groups) exist to create space for a critical discussion on assessment results. Nevertheless, some countries have planned for such mechanisms. The Gambia and Zambia have developed assessment policies with provisions relating to information flows. In The Gambia, the Coordinating Committee Meetings (CCM), held at least once a year in each of the country's six regions, are the most significant formal mechanism through which the NAT and EGRA results are shared and discussed among a large range of stakeholders. The platform was described as comprehensive, relevant, country-led, participatory, inclusive, and mutually complementary.

In Zambia, the 2017 NLAF considers information flows as a prerequisite for the use of learning assessment data. It underlines the importance of having 'two-way traffic of assessment information', to ensure all practitioners can provide feedback and implement interventions (Zambia, 2017: 10) (see *Figure* 7). In practice, actors exchange information during joint annual reviews and the Policy Implementation Technical Committee meetings, which are key moments of communication and coordination on learning assessment data.

4. The interplay of national actors: between miscommunication and misunderstanding

Figure 7. The flow of assessment data through all levels of the education system in Zambia

Source: Zambia, 2017: 10.

The different processes involved in the management of learning assessment data (e.g. test development and administration, data treatment and analysis, dissemination) are rarely inclusive or participatory. Currently, many tasks related to learning assessments are often insulated within assessment teams that rarely consult other actors, especially when it comes to assessment design. For instance, in Guinea, Senegal, Namibia, and Zambia, units responsible for the curriculum do not provide inputs for writing national assessment items. However, it is also important to mention that in order to participate in these activities, specific skills are needed.

In Guinea, Namibia, and Senegal, key actors in charge of assessments, planning, curriculum, or quality assurance only cooperate on an ad hoc basis, and therefore there are few opportunities for concerted work on assessments. Yet the existing body of literature confirms that the direct involvement of different actors in data management processes is likely to increase their awareness of assessment results and their importance, as well as informing them of the ways they can be used. More specifically, the participation of a wide range of education stakeholders in assessment design allows their concerns to be better taken into account, while

their involvement in data analysis increases the chances of the data being better understood and owned (UIS, 2018b). The SEACMEQ assessment strategy paid particular attention to this aspect. Policy issues were considered in the design of the assessment tools, to answer questions that policy-makers had, thereby linking policy concerns with assessment throughout its technical stages: 'One of the most important generalizable features of S[E]ACMEQ lies in the "policy thread" that runs throughout the entire research cycle. That is, the research cycle starts with policy concerns, moves through a highly systematic and focused applied research cycle, and then is completed with research-based suggestions about how to address the initial policy concerns' (Saito, 1999: 3).

Ad hoc exchanges do not translate further into meaningful dialogue on learning assessment data. One representative of Senegal's Ministry of National Education underlined this issue in the country: 'The central level must be well informed. Sharing is done within the ministry in important meetings, but there is no exploitation of assessment data or dialogue within the ministry and the education system.' A SABER report on Ghana student assessments observes a similar issue: 'Although there is general recognition that Ghana's results improved from TIMSS 2003 to TIMSS 2007, there has been little discussion in the country about how this happened' (World Bank, 2013: 11).

4.1.2 Looking for deep causes of poor actor collaboration

Much of the debate has focused on the institutional reasons for poor cooperation between actors, and while they are indeed important, as demonstrated in previous chapters, our study sought to explore other explanatory factors. We

4. The interplay of national actors: between miscommunication and misunderstanding

attempted to elaborate on this issue, seeking its roots in the MoEs' broader 'work culture' and actors' interpersonal dynamics. This aims to highlight the importance of actors' perceptions and behaviours, and it may account for why institutional solutions (e.g. a better definition of roles and responsibilities) only solved this bottleneck to a certain extent. However, it is also important to underline that it is very difficult to study such dynamics[26] and even more so to adopt measures to move towards a more positive equilibrium, as such contextual features are likely to evolve at a slow pace.

The quality of communication and cooperation on learning assessment data often reflects a broader 'atmosphere' or 'culture' characterized by certain practices in place within MoEs. Consequently, some environments are more conducive than others for smooth collaboration in general, and when it comes to learning assessments more specifically. For instance, the 'open-door culture' observed in The Gambia facilitates formal and informal exchanges. Some interviewees explained that it has become commonplace to call or meet their colleagues in their offices when they have requests, and to discuss them informally. This open communication culture is also reflected in the AU's relationship with the planning unit. The latter described their close relationship with the AU as very cordial: 'Relationship with the AU? They are our bosses, they can take out our staff at night. I cannot see a relationship that could be more cordial than this [everyone laughing]. That one is a public relationship.' This informal and

26 The study examined ideologies and values, as well as the functioning of informal institutions based on what interviewees reported in the interviews and what we observed during observations. It is therefore important to introduce a certain degree of caution, as the study mostly gathered perceptions of actors' cooperation dynamics, which might contain a certain bias. In addition, the study did not examine 'work culture' in depth, although some interview questions touched upon this issue.

cordial relationship facilitates collaboration between both teams, and The Gambia is a rare case where the planning unit is very much engaged in assessment data management. They are responsible for storing the data and using them for planning purposes, conducting a subject analysis for the NAT and presenting the results to stakeholders at the CCMs, sharing learning data with other actors in the system, and disseminating these data to schools in the form of scorecards.

By contrast, the division of structures appears clearly in the interviews and observations conducted in Guinea. For example, this is mirrored in the fact that, when asked about the intended use of learning assessments, multiple interviewees directed us back to the AU, saying this was their responsibility. According to the representative of the Institut Supérieur de Sciences de l'Éducation de Guinée (ISSEG), the lack of cooperation stems from the fact that people do not perceive themselves as being part of a 'systemic and complementary' whole. This is reflective of what UNESCO and UIS noted:

> The units which are responsible for these three [assessment, curriculum, and pedagogy] domains may not have a sense that they are responsible to each other or that they should be supporting each other. In some cases, this situation is compounded by the very organizational design of ministries of education, which may keep these units apart and not accountable to one another. (UIS, 2018b: 81)

4.1.3 Technical/political dichotomy?

For assessment results to trigger changes, AUs need to be able to make their voices heard and take these results into the political sphere. Our study suggests that, in many cases,

4. The interplay of national actors: between miscommunication and misunderstanding

how other actors perceive the status of AUs and the evidence they produce influences the use of learning data in planning and policy-making processes. Typically, these perceptions are those of a dichotomy between what should be *political* and what should be *technical*, which reduces the likelihood that learning data will inform decision-making processes, as illustrated below.

Divisions between technical and political profiles and the absence of bridges between them negatively affect the use of learning data. The impartiality of AUs is essential to avoid political interference in the assessment of student learning. However, in some cases, actors perceive that assessment teams' role should be completely disconnected from the decision-making process and limited only to technical aspects. Consequently, some assessment teams see themselves as not having a stake in how the data they produce are further used in the system, as was the case in Namibia and Guinea. One belief that often emerged from the interviews is that the political and technical dimensions are two very distinct aspects of policy-making, and therefore need to be handled by different structures with the corresponding mandates.

This may lead to AUs overstepping their responsibilities if they seek more influence in the policy-making sphere. This was particularly salient in Guinea, where some interviewee comments clearly reflect the dichotomy between those aspects perceived as falling under the technical and those perceived as political aspects of the policy-making process:

> *SNESE's mission is not to explain why, that's for INRAP and directorates like mine to see why. SNESE does assessments: 'these are the percentages.' But the refinement of the research is done*

4. The interplay of national actors: between miscommunication and misunderstanding

> *by INRAP.* (Representative of the Directorate for Basic Education, Direction Nationale de l'Enseignement Fondamental [DNEF])
>
> *The IT*[27] *unit produces a lot of information, and this is a technician's concern. Preparing the report is their job and then it is the responsibility of the decision-makers to interpret and use it'* (Representative of the department responsible for examinations, Service National des Examens Scolaires et Concours [SNESCO])

Linking assessment results with policy formulation and finding a balance between assessment teams' involvement in policy dialogue and their impartiality become a difficult exercise in this context.

Moreover, contexts where technical expertise in assessments is not widespread within ministries but concentrated in a narrow group of individuals can reinforce the perception that assessments are the exclusive domain of assessment teams and can discourage cooperation. As noted by one respondent in Guinea, a certain 'technical arrogance' could be observed towards colleagues less familiar with data. In addition, one respondent in Senegal reported that INEADE claimed that '[they] know how to do it' and was not therefore receptive to external assistance, even if it could have improved the quality of assessment data and reports.

27 This refers to the Statistics and Planning Department (Service des Statistiques et de la Planification).

4.2 The top-down dynamics between central and decentralized levels

4.2.1 The top-down nature of exchanges on learning assessment data

Although essential in the chain of assessment data management, decentralized-level officials are often relegated to different phases of the process. Their needs are frequently unaddressed by sample-based assessments, and feedback on assessment implementation and results is often weak or absent. Dynamics among central- and decentralized-level officials, often reflecting their top-down nature, lead to a lack of awareness, ownership, and use of assessment data at decentralized level.

As previously explained, the assessment design to a large extent determines who the final users of the data will be, as sample-based assessments cannot provide a picture of learning at the most local levels (e.g. district, school levels), unless they are purposefully created to do so. This is why it is important to clearly define the intended users of learning data at the inception of assessments and design them accordingly. If the intended use of assessment data is to inform decision-making at a macro level, subnational levels will naturally make only limited use of that data. However, subnational authorities need an objective measure of their students' performance to make decisions in their respective jurisdictions, and this should be considered in the design of assessment systems. While schools can rely on formative assessments, this is not the case for upper administrative levels.

4. The interplay of national actors: between miscommunication and misunderstanding

Our study demonstrates that learning data are not always representative at different administrative levels, which impedes their use. In Ghana, this is one of the main issues in mobilizing the NEA results for district-level officials. In Guinea, although one of the objectives of the *Évaluation Nationale des Acquis* is to monitor learning at decentralized level (UIS, 2015), the data are only representative at regional level and cannot offer information at district or school level.

The management of assessment data is mainly centrally driven. One of the rare opportunities for the decentralized level to participate in the process of assessment management is the test administration stage. To this end, regional or district officers meet central assessment teams in training or workshops aimed at preparing them for the supervision and implementation of the tests. However, many interviewees complained that once the data are returned to the central assessment team, they rarely receive written reports on the implementation and results of the assessments, which is a major demotivating factor. As described in *Chapter 3*, decentralized-level authorities are the most affected by ineffective dissemination channels. Officers from Guinea's Regional Inspections explained: 'When they come to assess, they come to do the work and leave. In return, we are not given any data. So we here are not informed. We don't have the data on learning … Even the reporting is not done.' A UNICEF representative in Namibia made a similar observation:

> *Often these assessment processes are designed and implemented in a top-down manner, and by the time they trickle down to school-level and community-level stakeholders, a lot of the information concerning the relevance of these assessments dissipates. All that users are left with is the trauma of assessment.*

4. The interplay of national actors: between miscommunication and misunderstanding

Consequently, in most countries, subnational officials showed greater awareness of examination data, which constitute their primary source of information on student learning. Few officials were aware of large-scale learning assessments, and many would naturally think of examinations when asked about assessments.

The study demonstrated that actors' knowledge of LSAs is, to a large extent, driven by their closeness to the ecosystem of learning assessments and opportunities to participate in their management. As decentralized-level officials are often far removed from the decision-making centres and barely involved in other steps of the process, aside from assessment implementation, there is hardly any opportunity for them to become more familiar with the assessment data, which ultimately decreases the likelihood of them being used, as pointed out by a civil society representative in Ghana: 'The further you go, the less data is used, sometimes not at all.'

The issues mentioned above are part of a broader context of centrally led education governance systems. Although decentralized-level officials play a fundamental role in the implementation of education policies, their role is often limited to implementing a centrally defined agenda. Verger, Novelli, and Altinyelken (2017) confirm this: 'The professionals who ultimately have to make new policies work (teachers, principals, local government officials, etc.) often perceive education reform as something imposed from above' (Verger, Novelli, and Altinyelken, 2017: 25). In Zambia, for example, respondents described relationships between the central and decentralized levels as 'disjointed'. A UNICEF representative in Zambia confirmed that while the primary use of learning assessment data is at national level, the lower levels usually

'just receive directives'. In addition, one respondent further explained that provincial officers were only invited to central level when there was noticeable poor performance in their area, which impedes more meaningful engagement of decentralized officials with assessment results.

4.2.2 Closing the feedback loop

Even in countries where learning data reach lower administrative levels and mechanisms are in place to provide feedback based on assessment results, these are not always effective, so that data are consistently and meaningfully used across administrative levels. In The Gambia, although learning data inform the preparation of SIPs, the weak monitoring of their implementation raised concerns about their effectiveness. In Ghana, although there are different spaces to discuss learning outcomes with the community, there are also doubts regarding the effectiveness of the system. The Education Sector Analysis (2018) notes:

> Anecdotal evidence suggests that parent-teacher associations and School Management Committees are largely inactive, thereby contributing to poor community oversight of school management and results. School Performance Improvement Plans are often absent from schools, and School Performance Appraisal Meetings are not held. Circuit supervisors often lack adequate training in school performance support and monitoring, and District Education Oversight Committees are often non-operational. (Ghana, 2018: 43)

In Namibia, there is no feedback loop as the central level does not provide support for monitoring the use of learning data at decentralized level. There is therefore little incentive for regional education offices to make use of the data, as there is

4. The interplay of national actors: between miscommunication and misunderstanding

no particular feedback expected, apart from the reporting on regional average student scores. The following quotes from representatives of the assessment team are illustrative of the weaknesses in the feedback loop:

> *In other words, we literally produce the data and send it through the regions, and it's up to the regions to see how best they can take it further, depending on their budgets or their way of doing things. We don't really have a stake in that.*

> *... we do not focus on the feedback; we leave it up to them to decide what they do with the information.*

A lack of communication between central and decentralized levels leads to mutual misunderstandings. As underpinned by interviews and observations conducted in Namibia, Senegal, and Zambia, interactions between central and decentralized levels can take the form of a 'blame game'. On one hand, decentralized-level officials often suffer from the top-down nature of their relationships with central offices and criticize them for not being sufficiently aware of local realities. On the other hand, multiple central-level officials showed dissatisfaction with the current use of learning assessment data by local authorities and sometimes attributed the ineffectiveness of educational reforms to their deficient implementation on the ground. This is illustrated by a quote from a Zambian representative from the national assessment team: 'We send [assessment results] to the schools, to the districts, to the provinces and to headquarters, but usually you find that the documents just gather dust.' In addition, in Namibia, the meeting for the validation of the 2019/20 annual plans at the ministry was an opportunity to observe the interactions between regional directorates and their colleagues at central level. Central directorates felt that their

decentralized level counterparts did not have the capacities or the understanding of how to develop annual plans that focus on strategic issues.

Subnational actors need support in different areas related to the use of learning assessment data. The observation made by Lai and Schildkamp (2016) about teachers' use of assessment data also holds for subnational levels, i.e. 'too much pressure to use data, combined with too little support, leads to less effective data use'. They need to be accompanied in building an 'assessment literacy', that is 'the knowledge and skills to enact the data analysis cycle (i.e. to analyse, interpret, and use different forms of data for decision-making), including an understanding of how assessment data is positioned within the data analysis cycle' (Lai and Schildkamp, 2016: 81). In The Gambia, a regional education director suggested that transferring the data analysis to regions could help increase their capacities, as contact points with planning responsibility should not be the only ones trained in data use. An assessment team member in Senegal made a similar recommendation, wishing to see more research-oriented offices at decentralized level to help develop tailor-made recommendations for subnational jurisdictions.

Subnational actors need to benefit from the assistance of skilled colleagues at central level as well as development partners. Assistance should be provided not only in terms of capacity development but also through the exchange of information and provision of adequate resources. In addition, greater involvement at different stages of the management of learning assessment data can persuade local actors that these data have similar value to other evidence they

usually refer to (e.g. examination results), and foster better understanding and ownership of assessments.

> A lack of communication between central and decentralized levels leads to mutual misunderstandings.

To conclude, our analysis demonstrates that institutional settings and interplay between actors both affect collaboration. This chapter demonstrated that although there are some opportunities to share information at central level, they remain ad hoc and insufficient to guarantee continuous exchanges on learning data in most countries. Moreover, the dichotomy between political and technical attributions of different teams may impede the emergence of national leadership in assessment systems. We have also seen that relationships between national and subnational authorities often reveal top-down dynamics, which can hinder the use of assessment data at decentralized level. This necessitates the establishment of an effective feedback loop throughout different administrative levels to improve the use of learning data.

Chapter 5. International partners' role and influence

The development of learning assessment systems is not a neutral technical exercise, but rather a highly political one (Verger, Novelli, and Altinyelken, 2017) with inherent power relationships. To better understand the limited use of learning assessment data, it is important to unveil underlying actors' power relationships and the challenges they might create for the use of assessment data. Whereas *Chapter 4* focused on these dynamics between national actors, this chapter looks at how their relationships with international partners may affect the use of learning data.

5.1 International partners play a key role in development of learning assessment systems

In all six countries studied, international partners played a key role in the development and implementation of large-scale learning assessments, as they provided crucial technical and financial support. This is often more broadly expanded to the establishment of information systems to produce quality education data, as demonstrated by partners' commitment to building stronger EMISs.

The Education 2030 Agenda emphasizes the importance of learning assessment systems in the international debate on education quality (Wagner et al., 2012). International agencies have also been increasingly interested in assessment data to inform their decisions on how best to support low-income countries, improve the learning outcomes of their children (UIS, 2018a), and monitor the impact of different policies and programmes on learning (Ramirez, 2018). The increasing

5. International partners' role and influence

demand for evidence of 'aid effectiveness fuelled a demand for more countries, particularly developing countries, to participate in assessments' (Lockheed, 2013: 178). A number of international organizations involved in education, such as UNESCO, GPE, and the World Bank, have advocated for the importance of large-scale learning assessments to address the global learning crisis: 'While opinions vary on how best to address the learning crisis, there is widespread agreement that data on learning outcomes will be key to finding solutions' (Johnson and Parrado, 2020: 1). Moreover, Kamens and McNeely (2010) argue that the global culture of testing that bolstered the spread of assessments worldwide would not have been integrated into national contexts without the presence of international and regional organizations as agents of mediation of and adaptation to global changes in individual countries.

Learning assessments, especially international ones, are tied to monitoring the SDG4 indicators, which set 'global metrics' and become an accountability mechanism influencing the way education systems define their goals, which further strengthens the position of international agencies promoting LSAs:

> ILSAs have become linked to the pursuit of SDG4 and are used as an accountability mechanism in relation to aid. ILSAs can provide a common framework for setting targets and evaluating progress, influencing the way we frame educational goals globally and reinforcing their value as policy instruments for measuring progress. Several international organizations have vested interests in sustaining and extending assessments, but tensions may also emerge between international organizations in relation to their roles and the use of ILSA data. (Addey and Sellar, 2019: 12)

In the countries studied, international partners have played a founding role and have been strong advocates for the use of LSA data. While commitment to Education for All goals spurred the development of learning indicators (Tobin, Nugroho, and Lietz, 2016), international partners often initiated a rollout of LSAs in countries. For example, in Zambia, the introduction of NAS for Grade 5 aimed to evaluate the investments made under EFA (Zambia, 2005). One interviewee from Guinea recognized that partners' focus on LSAs has been increasing over time. In the Ghanaian case, LSAs were often introduced through external projects. For instance, USAID and the Ghana Education Service (GES) forged the Ghana Partnership for Education, and agreed to conduct EGRA and EGMA (Ghana, 2014). Partners also participated very actively in the development of LSAs and supported all stages of assessment management. In Senegal, the first cycle of SNERS was sponsored by the World Bank, while RTI International, USAID, and the World Bank spearheaded the first cycle of EGRA. In Namibia, the introduction of NSAT was recommended by a World Bank study.

5.2 National authorities' autonomy in decisions restrained by dependence on external funding

Such strong influence from external partners cannot be fully understood without delving into the political economy underlying assessment systems. Benveniste (2002) underlines that 'assessment is also a political phenomenon that reflects the agendas, tensions, and nature of power relations between political actors' (Benveniste, 2002: 89). Kellaghan, Greaney, and Murray (2009) discuss the facets of assessment's role in the exercise of control and power in education. They argue

that assessments originate in a political process, often inspired and fashioned by political motivations. In addition, the form of assessment will be the result of competition between social actors. Moreover, an assessment can involve mechanisms for regulation and for holding social actors accountable, implicitly or explicitly, for outcomes.

Our evidence shows that the management of learning assessments is marked by a power imbalance, to varying degrees, between development partners and country decision-makers. Verger, Novelli, and Altinyelken (2017) argue that: 'International players have an increasing capacity to settle education agendas and define the priorities of countries concerning education reform processes through advice and technical assistance, but also to impose certain policies via funding and aid conditionality' (Verger, Novelli, and Altinyelken, 2017: 5). Our research confirms that power imbalances occur in different areas of the education sector in general, and more specifically in decisions concerning participation in LSAs and their modalities.

> "Learning assessments, especially international ones, are tied to monitoring the SDG4 indicators.

Firstly, because in most countries there is currently no adequate national budget for learning assessments, external funding and technical support remain essential for the implementation of learning assessments. When an assessment round is not sponsored by any international partner, countries face difficulties funding and implementing learning assessments, which sometimes leads to the cancellation of some cycles (as noted above in *Chapter 3, Section 3.5.2*). This issue is also highlighted in the literature:

5. International partners' role and influence

> Repercussions of externally financed assessments could account for the uneven and sporadic nature of assessments in many countries. More worrisome still is that externally funded assessments could contribute to the lack of alignment between national assessment and other components of the education system, which paves the way for limited national ownership of the assessment results. (UNESCO, 2019: 52)

However, as Addey and Sellar (2019) note: 'Non-participation may not be an option, even when ILSA data is not relevant to a country's specific educational or economic challenges' (Addey and Sellar, 2019: 11). The participation might be a requirement linked to membership of a given international organization, and non-participation might send a signal that the government is not committed to education and improving student outcomes (Addey and Sellar, 2019).

As the stakes associated with external partners' presence or withdrawal are high, they tend to increase their decision power, including when it comes to assessment design modalities. For instance, in a resource-constrained country like Guinea, where the education sector relies heavily on external funds, the weight of partners' contributions introduces asymmetries of power and the Ministry of National Education and Literacy has limited autonomy in terms of the choice of assessment modalities. The study demonstrated that this leads national officials to question the relevance of assessment data for their work and to stress the need for the ministry to regain control over assessment design. The importance of preserving the country's autonomy in the education sector in general has also been raised in Senegal. A number of respondents emphasized the influence that development partners have on policy-making, while underlining the significance of preserving their 'sovereignty', a concept that appeared

repeatedly in their interviews. Interestingly, 'sovereignty' was also mentioned by a World Bank representative but in quite different terms, showing that this has been a sensitive issue in the country: 'The government speaks about sovereignty, but there can be no sovereignty without resources'.

By contrast, in The Gambia and Zambia, relationships between national units and international actors were reported as balanced and beneficial, with back-and-forth dialogue between different parties. However, in The Gambia, the World Bank also plays a significant role in decisions related to assessments. Interviewees reported that it pushed for the introduction of EGRA. Although some country officials suggested that the assessment should be differentiated by grades, as is the case for the NAT, considering differences in curriculum, these changes were not adopted.

Another point of divergence between national and international partners typically concerns census-based vs sample-based assessments. While some ministries are in favour of annual census-based assessments (e.g. Ghana, The Gambia), international partners have often preferred sample-based assessments in view of their lower costs. In Ghana, a World Bank representative reported that when partners support assessments, they make their own calculations in terms of sample costs and the grade they want to evaluate, which explains why they disapprove of the shift towards a costly national census-based assessment. In The Gambia, partners recommended reducing the frequency of the NAT to lower the costs of the assessment, but the ministry kept the annual assessment.

5. International partners' role and influence

Many national actors expressed frustration at not being able to make their voice heard. A statement from a Ghanaian official working at decentralized level summarizes the challenges that national teams face:

> *We depend solely on donor support, and if the support is saying that 'this is the money I am bringing, and this is how I want you to use it', you will not say no. So you would oblige with that and suspend whatever you think will help you. So I think it is high time we also sit down and design our own learning assessment, tools, and use our own resources.*

5.3 Dependence on external capacities entails risks

Currently, scarce capacities and material resources (e.g. computers, software) do not allow all assessment-related tasks to be performed internally, since technically demanding tasks that ensure the validity of assessment instruments, sampling procedures, and analysis require highly trained staff, a requirement that falls short in the study countries. Consequently, different phases in the assessment management cycle (e.g. sampling, data analysis, report writing) are sometimes outsourced to international agencies, data contractors, or universities.

Outsourcing of some assessment activities can mitigate a lack of national capacities, ensure that the required standards in terms of data processing are met, and help avoid political influence (ACER-GEM; UIS, 2017). In addition, the Principles of Good Practices in Learning Assessment (ACER-GEM; UIS, 2017) explains that in some cases, outsourcing specific tasks can be more efficient or more cost-effective than conducting them internally.

5. International partners' role and influence

Nevertheless, outsourcing learning assessment activities contributes to restraining the influence and responsibility country experts have over different components of learning assessment systems, as well as on the use of the resulting data, particularly when they provide limited inputs to the development of assessment instruments (UNESCO, 2019). In these situations, organizations or institutions in charge of outsourced activities might be perceived as 'controlling' and as the 'owners' of the data (UNESCO, 2019). This is more likely to happen with cross-national assessments, since the lead agency and data contractors prepare country datasets in a standardized way before they are shared with national teams (UNESCO, 2019). Although it might be less of an issue with national assessments, their modalities can be largely influenced by partners as well, as confirmed by UNESCO (2019).

Even if outsourcing may solve some immediate concerns, resorting to this option too frequently should not deviate stakeholders from a longer-term commitment to sustainably build capacities, as the aim should be to reduce the need for outsourcing (UIS, 2018b). Additionally, while all aspects of assessment programmes should be open to outside scrutiny (ACER-GEM; UIS, 2017), it seems difficult to maintain a fully transparent and inclusive approach along the process of data management if in some cases certain tasks are performed in another country. For instance, EGRA reports in Ghana, Senegal, and Zambia have been prepared by RTI International. Quotes from Ghanaian national actors illustrate issues that might emerge when assessment activities are outsourced:

> *EGRA and EGMA were problematic: I would need to send an email to DC for clearance. We did not have easy access to that data, we needed to ask permission and send an email.* (MoE representative)

> *The release of the national EGRA/EGMA data needed to be approved by the USAID staff in Washington.* (World Bank representative)

> *It was analysed online using an application designed by the USAID board. I remember it was adopted from the state of Virginia. I don't have details of how it was done, but I know it was an online system that did it.* (Regional planning officer)

Outsourcing activities related to learning assessments entails a number of risks that can ultimately create the perception that international agencies are the 'owners' of the data (UNESCO, 2019).

5.4 Associating assessment data and results-based schemes

Learning assessment data are increasingly used in results-based funding schemes. The introduction of results-based financing in the GPE funding model[28] in 2015 is an example of such a trend. Literature specifies that performance-based accountability has become the prevailing form of accountability (see, for example, Verger, Parcerisa, and Fontdevila, 2019), and excessive use of learning data in M&E processes can lead to an over-emphasis on national accountability towards international partners. The case of Senegal is an interesting example, as the system has been openly geared towards performance-based management, due to development partners' influence. Learning data inform disbursement-linked indicators in strategic plans, development partners' projects, and performance-based contracts (see *Box 4* for more information).

28 The financing allocation is split between fixed and variable parts. A fixed part (70 per cent) is allocated based on certain requirements, among which is a commitment to developing national learning assessment systems, while a variable part (30 per cent) is disbursed based on results, i.e. when performance indicators demonstrate effective progress (GPE, 2015).

5. International partners' role and influence

BOX 4

The use of learning assessment data for disbursement-linked indicators in Senegal

The circular on general education policy in Senegal showcases a shift towards results-oriented and performance-based management of education (Senegal, 2018). This is further evidenced by the increasing utilization of disbursement-linked indicators, of which some are based on learning assessment data. For instance, the USAID-funded Partnership for the Improvement of Reading Skills in Primary Schools introduced institutional incentives through a system of school and district grants conditional on indicators of learning gains (USAID, 2013). Disbursement-linked indicators feature in the results frameworks of performance-based contracts, which bind actors with the lower administrative level, namely: (a) MoE inspections at regional level (Inspections d'Académie); (b) inspections at district level (Inspections d'Éducation et de Formation) and schools.

Other development partners' programmes, e.g. the AFD and GPE-funded programme (Programme d'Appui au Développement de l'Éducation au Sénégal [PADES[29]]) or ESP sub-programmes (e.g. Programme d'Amélioration de la Qualité et de l'Équité de l'Éducation de Base [PAQEEB]) also resort to this type of indicator. Grants are provided on a conditional basis, according to the achievement of targets measured by learning assessment data. While the mechanism ensures a more systematic uptake of learning assessment data, it tends to overemphasize their use for accountability purposes over other types of use throughout the planning cycle, such as policy formulation.

29 The Programme d'Appui au Développement de l'Éducation au Sénégal (PADES) is a four-year programme that supports the implementation of the sector plan PAQUET-EF (2018–2030).

5. International partners' role and influence

While such schemes may hold actors accountable, they also entail risks. In the 'payment by results' frameworks, a key assumption is that financial incentives will align donor and recipient goals (GEM, 2017). However, the *Global Education Monitoring Report* on accountability (GEM, 2017) documents that it may reduce national ownership over the policy changes that it intends to stimulate, raising doubts about the sustainable impact of such an approach. If several measures, including the development of learning assessment systems, are developed in the framework of results-based schemes, then the question arises of to what extent such measures are likely to be genuinely country-owned and to sustainably build an assessment data culture.

In addition, there must be an agreed definition of learning outcomes and indicators that are intended to measure them (GEM, 2017). However, assessment discontinuity and a lack of trust in the national assessment data can be a challenge for the measurement of target achievement. Donors may therefore prefer to set up parallel M&E systems instead of relying on national data, which is detrimental to the institutionalization of measurement systems (GEM, 2017).

Moreover, learning assessment results can have important repercussions in terms of funding, depending on whether or not they confirm the fulfilment of target objectives. Although not disbursing aid can prove difficult when targets are not achieved (GEM, 2017), the consistently low learning outcomes caused several donors supporting education, such as Irish Aid and FCDO (previously DFID), to suspend their investments in Zambia (GPE, 2019b). This not only adds to the unpredictability of funding flows (GEM, 2017) and challenges concerning the continuity of learning assessments, but it also consolidates

the pressure exercised by development partners. Finally, another potential consequence is that results-based schemes inform strategies exclusively aimed at affecting assessment results (UNESCO, 2019).

5.5 Strategic donor funding: learning assessment data to capture funding?

Low student performance in learning assessments is sometimes used as an important argument for capturing funding, as reported by some interviewees across countries such as Guinea, Senegal, Zambia, and, to a lesser extent, The Gambia. Our data reveal that learning assessment data play an important role in securing funds by 'proving that the needs are there' – not only for learning assessment systems but for the education sector in general; in other words, they legitimize external support. However, it is important to note that our research does not show that this constitutes the main motivation for participating in learning assessments. This is corroborated by Addey et al. (2017), who explain that 'empirical data suggests that low and middle income countries do not passively take part in [international large-scale assessments] to access funding without being driven by a more complex ensemble of rationales' (Addey et al., 2017: 7).

5.6 Weak capacities and national leadership in assessment systems leave governments unable to counterbalance partners' influence

Although development partners played a key role in the development of and support for learning assessment systems, they cannot be their sustainable driving force. National leadership in assessment management is essential

if learning assessment data are to spur policy action and bring about long-lasting effects on education systems. However, several conditions necessary for this leadership to emerge have not yet been met.

5.6.1 Since international actors cannot be a sustainable driving force for change…

Although direct budgetary support is now provided for assessment activities more often than before, partners' interventions remain largely project-based, which entails a number of risks. In some cases, partners conduct their own assessments for the purpose of specific programmes without drawing on the existing data. Such a situation was observed in Ghana, where FHI360 relied on the data they produced in the framework of the USAID-sponsored Learning Programme without consulting other assessment data available in the sector. Based on the analysis of East African countries, Elks (2016) recommends that '[d]onors could also consider how they can move towards using locally produced assessments to track programme performance, rather than administering their own assessments. This would allow national MoEs to set the appropriate standards for their system, reduce the expense of administering multiple assessments, and raise technical assessment capacity within the country and region' (Elks, 2016: 25).

Secondly, by definition, projects are time-bound. Therefore, the long-term sustainability of project initiatives depends on the government's capacity and interest in taking them over. Moreover, partners' programmes in the education sector often suffer from poor coordination, which also affects initiatives related to assessments. Focusing on some specific

areas of intervention (e.g. improvement of early literacy) can result in overlaps between a multiplicity of stand-alone programmes. This issue was mentioned in Senegal, where many respondents observed poor donor coordination and the absence of a framework leading to regular overlaps in their projects. The PADES presentation document confirms that '[m]any international partners intervene in the sector in support of the sector programme, mainly with a project-oriented approach. Ultimately, the fragmentation of aid is important for the effectiveness of interventions, but ad hoc and limited in time, at the expense of the sustainability and harmonization of the MEN's actions' (Senegal and AFD, 2018). Consequently, the ministry has difficulties asserting itself and establishing a framework to coordinate and harmonize development partners' initiatives. A similar duplication of effort has been observed in The Gambia regarding multiple reading programmes. However, the issue was resolved with the harmonization of three different programmes for teaching early grade reading (namely Jolly Phonics, Serholt Early Grade Reading Abilities, and National Languages) into The Gambia Reads programme in 2015.

While partners' initiatives may overlap in certain policies and education levels, some education levels or geographic areas may remain marginalized, such as secondary education that is only targeted at a few learning assessments (see *Chapter 2, Section 2.3*). Moreover, in Guinea and Senegal, targeted interventions in specific geographic areas have led to the creation of a 'two-speed education system' in which the schools benefiting from donors' support are better off.

5.6.2 ... there is a need for national leadership in assessment systems

For the reasons mentioned above, there is a need for strong national leadership in assessment systems. National leadership is core to making sure assessments address countries' data needs. It is also essential to sustainably developing assessment systems and to bringing different actors working on assessments together: 'Leadership is required to achieve broad acceptance of learning assessment among all key stakeholder groups concerned ... The government's strong support for the assessment system is essential to engage stakeholders effectively in the assessment system, ensuring that contributions are directed towards the improvement of learning' (GPE and ACER, 2019: 5). Nevertheless, limited expertise and resources, as well as institutional issues, pave the way for difficulties that impede the emergence of national leadership.

Distrust in the AUs' capacities and the reliability of the data they produce, as compared to that produced by external partners, may be a barrier to the emergence of national leadership. In addition, pre-existing institutional challenges are not conducive to the emergence of this leadership either. The absence of 'unity' (i.e. lack of a common understanding and shared appreciation of learning data), as well as smooth cooperation between actors in charge of assessments, creates further obstacles. In such contexts, external partners may act as local mediators. For instance, in Namibia, UNICEF created, at the ministry's request, a steering committee led by NIED, Programmes and Quality Assurance (PQA), and DNEA, aimed at conducting research in aspects related to the learning assessment system.

5. International partners' role and influence

Competing views and aspirations between assessment teams and other units in charge of assessments might impede the emergence of leadership in learning assessment systems. If the relationship between different units in charge of assessments is that of competition rather than collaboration, it creates obstacles impeding the emergence of national leadership. Guinea is an interesting example: the major obstacle to the emergence of national leadership is linked to doubts regarding the authority of the assessment team. Competition between key units in charge of assessments (i.e. SNESE and INRAP), tensions between SNESE and other ministry of national education and literacy units, as well as uncertainty regarding SNESE's institutional status and its limited decision-making power, all contribute to this issue. In the midst of this institutional turmoil, development partners managed to build a partnership with SNESE, but in this context, other actors in the system perceive that evaluations undertaken by the assessment team mostly emanate from development partners rather than the ministry itself. In addition to the implementation and management of regular LSAs, partners regularly request and sponsor studies related to their fields of intervention (e.g. a study on the impact of multi-grade classes sponsored by UNICEF). This situation may have created an excessive reliance on partners who actively participate in the definition of the assessment team's objectives and missions. One UNICEF representative explained: 'Partners have a lot of influence, but SNESE has a certain margin to give its opinion and influence things. But given the capacities of some people or their concerns, it doesn't happen this way. They simply wait for partners to define their assessments.'

A clear high-level leadership in learning assessment systems facilitates collaboration. Our study demonstrated that the existence of a unit able to bring different teams working on assessments together and oversee their work is an important factor for improving collaboration and the use of learning data. For instance, the ECZ in Zambia has become a key institution in the national assessment system, as it plays an important role in coordination. In The Gambia, the AU is the focal point for learning assessments and it brings together different actors working on assessments. The Assessment Policy (2015–2022) provides it with the mandate to 'coordinate the design and administration of these assessments as well as the management and dissemination of the results' (Gambia, 2015: 25).

> "limited expertise and resources, as well as institutional issues, pave the way for difficulties that impede the emergence of national leadership.

5.7 Learning assessment data need to be owned to be used

One direct consequence of international organizations' control over learning assessment design is the limited national ownership of assessments, leading to the misalignment of data production and data use:

> In many countries, the lack of alignment between data production and data use is also a direct consequence of the limited ownership of assessment data. It is feasible to venture that, if countries do not sense that they hold some power over the assessment process, then they will feel less inclined to use the data and act upon the evidence. This lack of ownership is in turn the result of countries' limited participation and influence

> on different parts of the assessment process, including funding, design, data management and the dissemination of results (UNESCO, 2019: 52).

In all countries, international partners are considered heavy users of learning assessment data. Learning assessment data serve to inform country strategies and gauge the efficacy of their interventions. For example, EGRA measures the impact of partner-supported reading projects in a number of countries (i.e. Ghana, Guinea, and Senegal). In Guinea, UNICEF relied on learning assessment data as part of its projects aimed at promoting girls' education, while in Zambia they have been used to develop the 'Let's Read Zambia' Campaign, a USAID-based literacy programme targeting all primary schools in selected remote provinces.

Although it is essential for partners to rely on data from learning assessments to inform their activities, the study shows that the use of these data between international and national actors is unbalanced, with the former often making greater use of it than the latter. For instance, several respondents in Zambia, Guinea, and Senegal stated that development partners were the primary recipients of learning assessment data. Deploring the limited uptake of assessment data by ministry directorates, one representative of the Guinean assessment team said that international partners were 'the real users' of their results.

Although most national officials that participated in the study were supportive of the use of learning assessment data, some of them perceived assessments as a 'procedure' with which they had to comply to satisfy donors' requirements, rather than a useful tool for their work. This perception was stronger

among officials at decentralized level and those that might be involved in assessment implementation but who do not receive appropriate feedback on students' results. However, '[a]ssessment merely to indicate a government's commitment to achieving Education for All goals, or to move into line with international expectations, is unlikely to result in findings being given serious consideration in policy-making, revised resource allocations, or being applied by managers to improve the education system' (DFID, 2011: 7).

A lack of national ownership of learning data has been documented in a number of countries. In Guinea, some respondents noted that the analysis of learning data ended up losing relevance for the ministry's work. One of the reasons that contribute to this is the lack of control that the ministry has over the design of a national assessment. One respondent indicated: 'We are going to do it because the donors are asking for it, the initiative does not come from the government. They are going to do the reports because they are asked to do them, not because they need them.' The issue of relevance of some assessments and their lack of alignment with the country's priorities was also raised in Senegal with regard to international assessments. As one assessment team member in Senegal put it, 'it is more relevant to work at decentralized level instead of taking time to make international assessments and write reports. Priority should be given to decentralized national assessments.' A statement by a Namibian decentralized official also clearly reflects the same issue: 'We use them very little, but we live with them.'

Lessons can be drawn from the Gambian case. Greater engagement and ownership of learning assessments in the country can be partly explained by the internal push

to develop the system, together with a strong commitment to building national capacities. A more proactive approach to donor funding (i.e. using assessment data to seek funding for country initiatives), facilitated by the active dissemination of assessment information to NGOs and international agencies, have helped to further anchor the latter's support and establish strategic partnerships. Instead of the asymmetrical nature observed in other countries, relationships between national and international actors seem to be mutually constructive. As noted by an M&E unit representative, 'the partners are there to give a guide and not to dictate what they think is relevant'.

5.8 Building a culture of evidence can foster the use of assessment data

> "the use of these data between international and national actors is unbalanced, with the former often making greater use of it than the latter.

It is important to go back to the very origin of learning assessments to fully comprehend the issues of ownership that can emerge. As instruments that are imported and exogenously influenced, they hold intrinsic cultural norms and values that may necessitate changes in recipient countries (Addey et al., 2017). The literature has shown that the use of assessment data is embedded in a more general organizational culture (Raudonytė, 2019). Although more or less effective assessment systems are in place, the use of learning assessment data is not yet rooted in policy-making practices at country level.

5. International partners' role and influence

Importing policies from other education systems with radically different features is not a neutral exercise, as they convey underlying values that might not be appropriate in the adopting countries and break with the practices in place. Verger, Novelli, and Altinyelken (2017) make an interesting contribution to this question:

> As the gap between the new policy and the previous system becomes bigger, implementation processes become more problematic (Rizvi and Lingard, 2010). This 'gap' is usually accentuated in relation to policies imported from elsewhere and initially designed by officials unconnected to local realities. (Verger, Novelli, and Altinyelken, 2017: 25)

While they are part of global discourse on evidence-informed decision-making, learning assessments introduce new practices and views on the management of education systems which need to find a favourable environment to be adopted in specific countries. We refer here to the beliefs and practices that feature assessment data as an essential component of decision-making and, more generally speaking, of actors' practices. Drawing on Rizvi and Lingard (2010), it is possible that when learning assessment data are perceived as too sharply contrasting with previous systems, especially when they are prescribed by external actors, there are greater chances of difficulties emerging in their use. Although in all countries there is an overall understanding that evidence and learning data are key to different planning activities and policy-making, there are mixed views about the current existence of a widespread culture of evidence within ministries.

5. International partners' role and influence

However, respondents have noted changes in actors' perceptions and attitudes towards assessment results over time. To some extent, countries have overcome political fears around publishing disappointing learning assessment results and have embraced more transparency. Numerous examples of governments holding assessment results back for political reasons have been documented in the literature, e.g. 1995 TIMSS in Mexico (World Bank, 2018). The political sensitivity of publishing assessment findings has also pushed some developing countries to stop participating in assessments; Botswana decided not to be involved in the Progress in International Reading Literacy Study from 1991 to 2011, and South Africa in TIMSS in 2007, due to the outrage that resulted from students' low performance (Raudonytė, 2019). Similarly, our evidence shows that, in Ghana, learning assessments used to be highly politicized. When learning assessments were introduced, the MoE often had a strong defensive reaction to poor results and did not always publish them, for political reasons. According to respondents, this also prompted the end of Ghana's participation in TIMSS. In addition, the methodology of assessments was frequently attacked, in an attempt to shift the attention from poor results. However, the situation has been gradually changing and the ministry is now much more open in this respect. According to USAID (2014), 'Zambia's initial EGRA/EGMA/SSME [*Snapshot of School Management Effectiveness*] results were so devastating that education managers were at first hesitant to deal with them publicly' (USAID, 2014: 11). The same publication explains that, after a change in government, increased media coverage was intended to build support for new education policies.

5. International partners' role and influence

To conclude, this chapter has demonstrated that although international partners have played a key role in the introduction of and support for learning assessments in the countries studied, their support had different impacts, depending on contextual in-country factors. Although in some cases their support resulted in mutually beneficial relationships that strengthened assessment systems and the use of their data, in others their influence was so strong that it overstepped countries' autonomy in key decisions on assessments' design and management. In some countries, this led to limited national ownership of assessment data.

Conclusions and recommendations

Understanding what students learn and what might influence teaching and learning processes is essential to improving education systems. Educational planners and other stakeholders throughout the education system (e.g. decentralized-level officials, curriculum and teachers' professional development units) systematically need to know: What are students learning? Who in the system is underperforming? What are the characteristics of those student groups? What are the main obstacles to learning? What are the priorities and strategies that can effectively overcome these obstacles? A multitude of LSAs have been put in place to trace student results over time and to try to answer these questions. However, unless policy-makers and planners actively consider, analyse, and engage with assessment results during the course of policy-making, these results lose their relevance. Reporting general scores, reflecting averages of student performance, is an important first step for diagnosing the overall performance of the education system. Nevertheless, assessment results have a much higher potential to inform education systems at all levels, while also providing insights for policy design, implementation, and evaluation. Fulfilling this potential is important, to help address a pressing need for sound education policies for improving student learning in sub-Saharan Africa.

It is also very important to emphasize that assessment data indicate aspects in need of attention, but they do not provide ready-made answers on how to improve learning. There is a need to evaluate the education system beyond student testing (e.g. teachers, institutional structures), using existing

administrative data and other available studies. In addition, learning assessments focus on cognitive foundational skills, often passing over other skills (e.g. socioemotional). Relying excessively on learning assessment results creates a risk of narrowing down education system goals to improving a set of limited indicators. While they can provide useful information on student performance in certain areas, they cannot be equated to the purposes of education systems (Breakspear, 2014). Defining the end-goals of education requires broader democratic deliberation: 'the discussion of educational end-goals involves ethical deliberation about what matters in education and what an educated person should be' (Breakspear, 2014: 11). Although we are not developing these issues further here, our findings need to be interpreted while bearing these considerations in mind. Only by fully understanding the risks of using data from LSAs can we reap their potential benefits.

However, processes for the use of evidence, or learning data more specifically, are not straightforward. Defining and researching how actors engage with evidence is intrinsically difficult. Policies and planning activities take many things into account, evidence being just one part of a complex puzzle. As argued by Nevo and Slonim-Nevo (2011: 1178), professional judgement is just as important a variable as others. The authors further argue that it is important to leave room for practitioners' constructive and imaginative judgement and knowledge, and that 'empirical evidence is better regarded as one component in the mutual and constantly changing journey of client and practitioner. … a wide range of information sources, empirical findings, case studies, clinical narratives and experiences are to be used in a creative and discriminating way throughout the intervention

process'. Finally, it is important to acknowledge the crucial role that politics play in the decision-making processes. Before drawing more general conclusions from the study, it is key to consider them in this broader context.

Our study has documented the use of various LSAs as well as barriers and conditions facilitating their use. Although dynamics that underlie the use of learning data vary in the countries of the study, our findings lead us to several reflections and guiding principles to consider when looking for ways to improve the use of learning data.

Ministry of education leadership

Clearly define the goals of your assessments

Learning data can be used for a number of different purposes (e.g. diagnosing the main issues in the education sector, informing macro-level policies, supporting the development of specific strategies, monitoring and evaluating the effectiveness of implemented policies, providing information to schools and subnational level officials). Nevertheless, one assessment cannot achieve all these goals and cater for the needs of all actors. In addition, assessment results cannot be retrospectively used for different purposes than initially foreseen. It is therefore important to have a clear idea of what kind of information assessments need to collect and for what purpose. This is not an easy exercise, as actors in the system have different needs and these need to be thoroughly evaluated to see how assessments may address them. The intended use of assessment data needs to be defined in a participatory process and then widely disseminated in related documents. The development of a national assessment

policy may help bring different actors together and critically reflect on the main goals of assessments and how they could coordinate to provide the most relevant information to different actors. It can also help provide a clear vision and strategic planning for the assessment system, which has the potential to improve the use of learning data.

Gear the system towards national leadership and ownership

Unless they are perceived as being part of the national data system that the government controls, large-scale learning assessments cannot effectively inform national planning and policy processes. National leadership and ownership of the management of these assessments are key to making sure national actors are involved with assessments and feel autonomous when defining their modalities, ultimately sustainably improving the use of data.

Plan for the use of data when developing assessments

For countries that are yet to develop their national LSAs or join a regional/international assessment, it is very important to make sure they plan for the use of assessment results early on. The availability of good quality assessment data will not automatically translate into effective use of that data. It is therefore essential to plan for dissemination and collaboration channels, communicate effectively on assessments' objectives, support actors in terms of capacity development for data analysis and interpretation, and establish effective feedback loops.

Commit to the development of national capacities in the production and use of learning data at different administrative levels

The study has clearly highlighted the central role that the development of national capacities plays in strengthening the use of learning data throughout administrative levels. Countries that focused on creating strong national assessment teams that are autonomous in their activities and trusted by other actors in the system, instead of outsourcing more difficult assessment tasks, managed to create greater ownership of assessment results and wider opportunities for the use of results. Training on the use of data for other actors in the system has often been neglected and it is currently needed, to ensure a more effective uptake of results.

Reflect on the best institutional position of your assessment team

Bringing assessments institutionally closer to units making decisions in different education areas, especially curriculum and planning, might increase the use of assessment data for those purposes but it disperses technical expertise on assessments, as different assessment teams often do not collaborate in an effective way. The closeness of the assessment system to the MoE and the concentration of assessment activities are important elements to reflect on when making these decisions.

National assessment teams

Based on a national vision for assessments defined collectively, choose assessment design options that respond to identified actors' needs

This study identified implications related to conducting different types of assessments (e.g. EGRA/EGMA, national, regional, or international LSAs) and choosing different assessment design options (e.g. sample vs census-based assessment). Choosing from among these options depends on the goals of a national assessment system and the goals of a specific assessment; it is therefore not possible to recommend one over another. Technical teams will be best placed to choose modalities that allow information to be collected that responds to the needs previously identified by participatory consultations.

Facilitate the understanding of learning data for different actors by adapting dissemination products and their content

Planning dissemination methods early in the programme while ensuring some flexibility later on is very important not only for planning financial resources accordingly, but also for ensuring the timeliness of dissemination. Considering a mix of dissemination products and methods, going beyond assessment reports and dissemination meetings, tailored to different groups of stakeholders, can help them make sense of this information for their activities. Concrete and realistic lines of action also need to be clearly identified for different actors. Evaluating dissemination effectiveness and the use of learning data by end users can support the improvements of future assessment cycles.

Reporting is not the end of an assessment cycle: link dissemination activities with constructive feedback loops

Throughout the study, it became clear that so much effort goes into the development, implementation, and reporting on assessment results that the cycle often stops there; assessment teams frequently have no more energy or financial and human resources to engage further with results. However, the real potential impact of assessments on education systems can be harvested only if reporting goes hand in hand with effective dissemination channels and follow-up activities on results. The work of assessment teams needs to extend to data dissemination and support to other actors working with these data. It can take the form of a feedback loop extending from the central level to regions/provinces, districts, and schools on specific findings that deserve attention. This needs to go hand in hand with making sure users understand concrete ways in which learning data can be used (e.g. developing guidelines/protocols to support teachers, education officers, and policy-makers) and have the capacities to engage in these processes meaningfully (e.g. conducting related training). Assessment teams also need support to undertake these new responsibilities, and their human resources require strengthening.

Involve other actors in the management of assessments

Involving actors in different assessment management steps is important, to gain their 'buy-in'. Although all actors cannot be involved in every step, as some of the assessment tasks are highly technical, reporting closely on those activities might help ensure actors remain informed. As one interviewee in Senegal rightly put it: 'If you have contributed to something,

there is a greater chance that you will use it.' It is particularly counterproductive to involve officials in only one step of data management and then not communicate further. The study illustrated that this often happens with decentralized-level officials who are involved in implementing assessments but disconnected from other data management activities, including receiving final reports. This creates negative feelings towards the assessment, which is seen as a procedure rather than a tool for informing their work. As they will be able to have a say and provide more inputs into the process, stakeholders will be encouraged to own and thus effectively engage with assessment data (UNESCO, 2019).

Look for synergies with other information sources, especially examinations

Learning assessment data are more informative if interpreted together with other evidence and cross-checked with other information sources. Examinations are widely used by different actors as a key reference for student learning outcomes. It is therefore important to see how synergies can be found between different assessments and how their results can complement each other (e.g. experimenting with such analysis in assessment reports).

Adjust assessment cycles to regular planning and budgeting activities

One of the difficulties that we observed in the study was mobilizing assessment data for planning when assessment results are not released in a timely way, or when their cycle is not in line with planning processes. It is therefore key to make sure that, if you have yearly assessments, they come in

when yearly action and budget plans are being prepared. If assessments take place every couple of years, make sure the data will be available for ESA and ESP preparation.

International partners

Invest in capacities and transfer expertise; discourage outsourcing

Partners supporting assessment systems in sub-Saharan Africa should ensure that capacity development is central to their activities by limiting outsourcing of assessments as much as possible. Support for developing national strategies for assessments is another potentially strong leverage for fostering national leadership.

Ensure that your support for learning assessments is in line with priorities defined in ESPs and other national strategic documents

The alignment with national priorities when supporting assessment systems should be a guiding principle in partners' activities.

Rely as much as possible on existing national assessments, rather than creating new ones

Conducting new assessments for specific projects creates a parallel assessment system, which disperses the human and financial resources of parties working on assessments. It is therefore important to rely on and strengthen national assessments and capacities.

Actors' dynamics matter

Although it might be difficult to act on barriers related to actors' interests and power dynamics, it is important to acknowledge that these factors are at play. Even if often we cannot target them directly, other recommendations provided in this book can help better understand the processes for the use of data and act on those processes under ministries' direct control (e.g. capacity development, collaboration, and dissemination channels).

Finally, processes around the use of evidence, and learning data more specifically, take time to develop, as they are often linked to a broader culture of evidence and its management. Learning assessments introduce new practices into the management of education systems and need to find a favourable environment, which takes time to build.

Annex 1. Glossary

Accountability: 'Accountability is a process, aimed at helping actors meet responsibilities and reach goals. Individuals or institutions are obliged, on the basis of a legal, political, social or moral justification, to provide an account of how they met clearly defined responsibilities' (GEM, 2017: xii).

Assessment design: 'The implementation plan for the whole assessment, including its purpose, the target population, the content to be tested, testing cycles, etc.' (ACER-GEM; UIS, 2017: 52).

Assessment framework: 'A document that underpins the validity of the assessment by making explicit the aim of the assessment, and what it covers in terms of content, skills, knowledge, and context. The assessment framework defines terms relating to the assessment, which means that when people discuss the assessment, they can communicate its purpose and characteristics clearly' (GPE and ACER, 2019: 53).

Assessment instruments: 'Test instruments and contextual instruments used in an assessment' (GPE and ACER, 2019: 54).

Assessment system: This 'is a group of policies, structures, practices, and tools for generating and using information on student learning and achievement. Effective assessment systems are those that provide information of sufficient quality and quantity to meet stakeholder information and decision-making needs in support of improved education quality and student learning outcomes' (Clarke, 2012: 1). The national assessment system encompasses all different assessment types that exist in a given country (IIEP-UNESCO, 2019a).

Census-based assessment: An assessment conducted based on the census, which is 'an official survey involving the whole population within a defined system' (GPE and ACER, 2019: 54). 'A census-based assessment can provide diagnostic data on

Glossary

each school, help plan interventions, and inform communities/parents about individual school performance. But it can also create perverse incentives, significantly increase costs and drive distortions to the teaching/learning process' (DFID, 2011: 3).

Cost-efficiency analysis: 'Comparing the cost of different inputs with their estimated correlations with learning outcomes' (IIEP-UNESCO, World Bank and UNICEF, 2014: 170).

Curriculum: 'A document that outlines what students are expected to learn in different subject or topic areas at different grade and/or age levels' (GPE and ACER, 2019: 55).

Decentralized level: Regional and local authorities exercising functions and power in the context of decentralization, i.e. the 'dispersion or distribution of functions and powers from a central authority to a local authority or community' ('UNESCO Thesaurus', n.d.).

Disaggregation (data): 'Disaggregation is the breakdown of observations, usually within a common branch of a hierarchy, to a more detailed level to that at which detailed observations are taken' (OECD, n.d.).

Disbursement-linked indicators: Indicators whose achievement triggers the disbursement of funding from development partners as per pre-agreed results (ADB, 2015). They 'can be classified along three main categories: 1) determining the level of results (outputs, outcomes and impacts or the inputs and activities necessary to strengthen programme performance); 2) measuring change in a direct or indirect way; 3) measuring change in a qualitative or quantitative manner' (GPE, n.d.).

Dissemination strategy: 'A dissemination strategy is developed to identify key stakeholders and their information

Glossary

needs, and the different dissemination products to address these' (GPE and ACER, 2019: 56).

Early childhood care and education (ECCE): 'Programmes that, in addition to providing children with care, offer a structured and purposeful set of learning activities either in a formal institution (pre-primary or ISCED 0), or as part of a non-formal child development programme. ECCE programmes are normally designed for children from age 3 and include organized learning activities that constitute, on average, the equivalent of at least 2 hours per day and 100 days per year' (GEM Report Team, 2007: 20).

Early childhood education (ECE): 'Early childhood education provides learning and educational activities with a holistic approach to support children's early cognitive, physical, social and emotional development and introduce young children to organized instruction outside of the family context to develop some of the skills needed for academic readiness and to prepare them for entry into primary education' (UIS, 2012: 79).

Education Management Information System (EMIS): An EMIS can be defined as 'a system for the collection, integration, processing, maintenance and dissemination of data and information to support decision-making, policy analysis and formulation, planning, monitoring and management at all levels of an education system. It is a system of people, technology, models, methods, processes, procedures, rules and regulations that function together to provide education leaders, decision-makers and managers at all levels with a comprehensive, integrated set of relevant, reliable, unambiguous and timely data and information to support them in completion of their responsibilities' (GEM Report Team, 2008: 101).

Glossary

Education sector analysis: 'Education sector analysis (ESA) is the first step in sector planning, and consists of conducting an in-depth and holistic diagnosis of recent trends and of the current status of the education system, to identify progress achieved and outstanding challenges' (IIEP Learning Portal, https://learningportal.iiep.unesco.org/en/issue-briefs/plan-for-learning/education-sector-analysis).

Education sector plan (ESP): A national policy instrument, developed under the responsibility of government, that provides a long-term vision for the country's education system, and outlines a coherent set of practicable strategies to reach its objectives and overcome existing challenges. It is based on a sound analysis of the current situation and of the causes of successes achieved and difficulties encountered. It should include implementation and monitoring and evaluation (M&E) frameworks (IIEP-UNESCO and GPE, 2015).

Education sector plan preparation: An iterative process that consists of the following interrelated stages (some of them are also part of a broader planning cycle) (IIEP-UNESCO and GPE, 2015):

1. Education sector analysis
2. Policy formulation and choice of strategies
3. Programme design
4. Action plan
5. Costing and financing
6. Implementation arrangements
7. Monitoring and evaluation framework

Educational planning: 'The application of rational, systematic analysis to the process of educational development with the aim of making education more effective and efficient in responding to the needs and goals of its students and society' (Coombs, 1970: 14).

Glossary

Equity: 'In education, the extent to which access and opportunities for children and adults are just and fair. This implies reduction of disparities based on gender, poverty, residence, ethnicity, language, and other characteristics' (GEM Report Team, 2008: 391).

Examinations: Examinations aim to certify the accomplishment of individual students and/or to select individuals into the next education level (Lockheed, 2016). Whether school-based or externally administered, they are often standardized in nature so as students are given an equal opportunity to show what they know and can do in relation to an official curriculum or other identified body of knowledge and skills (Clarke, 2012; Lockheed, 2016).

Implementation plan: 'Outlines the detailed activities for a specific period of the plan, with information on timing, roles, responsibilities, and costs' (IIEP-UNESCO and GPE, 2015: 23).

Inference: '[i]n statistics, the process of drawing conclusions about a parameter one is seeking to measure or estimate' (Britannica, 2016).

In-service teacher training: 'The process by which teachers engage in further education or training to refresh or upgrade their professional knowledge, skills and practices in the course of their employment' (TTF and UNESCO, 2019: 121).

International assessments: Assessments that 'provide information that is similar to that of the national assessment, but for more than one national education system. It is generally not sensitive to individual systems since its main goal is the comparability of the results among the participating countries' ('UIS Glossary', n.d.).

Large-scale assessments (LSAs) of students' learning: Assessments that can be sample-based or conducted as a census, in both ways representing a larger school population.

Glossary

They can be international, regional, or national in scope and focus on a particular population. In addition, LSAs: '(1) Are standardized to enable comparability across students, schools and in some cases, countries ...; (2) Are equally likely to be conducted in centralized or decentralized education systems; (3) In some instances can compare education systems across countries in the same region or internationally; (4) Do not have as their main purpose to certify individual student achievement, and do not refer to assessments used by teachers in classrooms, or to selective or "gate-keeping" assessments such as graduation examinations or university entrance examinations' (Tobin et al., 2015: 2).

Learning assessment data: Assessment data 'are information gathered through the learning assessment survey (questionnaire) that usually includes results from assessments of learning outcomes combined with the background information on student, family, community, school and teacher characteristics that was collected at the same period of time. It can refer to both reported results and raw databases. Findings from learning assessments could then be cross-checked with other evidence to inform policy-making' (IIEP-UNESCO, 2019a: 79).

Learning outcomes: 'Totality of information, knowledge, understanding, attitudes, values, skills, competencies or behaviours an individual is expected to master upon successful completion of an educational programme' ('UIS Glossary', n.d.).

Learning poverty: Learning poverty 'means being unable to read and understand a simple text by age 10. This indicator brings together schooling and learning indicators: it begins with the share of children who haven't achieved minimum reading proficiency (as measured in schools) and is adjusted

Glossary

by the proportion of children who are out of school (and are assumed not able to read proficiently)' (World Bank, 2019a: 6).

Monitoring and evaluation (M&E): Monitoring 'includes regular monitoring by the ministry of education under the plan's monitoring framework, and periodic joint sector reviews to assess progress and discuss solutions to bottlenecks. Evaluations usually take place at mid-term and the end of an ESP period to assess impact and outcomes, relevance, cost effectiveness and sustainability of strategies' (GPE, 2022).

Multivariate analysis: is the statistical study of data where multiple measurements are made on each experimental unit and where the relationships between multivariate measurements and their structure are important.' ('Multivariate Regression Analysis: Stata data analysis examples', n.d.)

National assessments: These assessments are designed to provide evidence about the quality of student learning outcomes in identified curriculum areas. In certain cases, they are administered to a sample of individuals. Some background information that is important for linking analysis to policy questions at the national, subnational, and local levels is also often collected (Postlethwaite and Kellaghan, 2008; 'UIS Glossary').

Outsourcing: 'Contracting an individual or agency located outside of the assessment agency to perform specific tasks' (ACER-GEM; UIS, 2017: 56).

Performance indicator: 'A variable that allows the verification of changes in the development intervention or shows results relative to what was planned' (OECD, 2002: 29).

Planning cycle: This refers to a larger cycle that is not equated only with the education sector plan preparation.

Glossary

It encompasses the following stages:[30] 1. Education sector analysis: analysing the current situation in the sector; 2. Policy decision: selecting overall goals and strategies; 3. Education sector plan preparation: defining precise objectives and programmes; 4. Implementation: annual planning through operational plans and budgets' updates; 5. Monitoring and evaluation: measuring progress for corrective action (IIEP-UNESCO, 2019a).

Political economy: 'A political economy approach generally aims to understand how diverging stakeholders' interests and incentives, which are themselves a result of a broader institutional, social and political context in the country, determine policy change. This approach therefore recognizes that "education reform does not take place in a vacuum, but under specific constraints and opportunities, many of which are politically driven, shaped by the interests and incentives facing different stakeholders, the direct and indirect pressures exerted by these stakeholders, and by formal and informal institutions. Each of these factors influences different aspects of education reform, whether policy design, financing, implementation or evaluation"' (Kingdon et al., 2014: 5)' (IIEP-UNESCO, 2019a: 92).

Pre-service teacher training: 'Teacher education before entering a classroom or other educational site as a fully responsible teacher' (TTF and UNESCO, 2019: 121).

Quality assurance: 'Quality assurance encompasses any activity that is concerned with assessing and improving the merit or the worth of a development intervention or its compliance with given standards' (OECD, 2002: 31).

30 For conceptualization purposes, a simplified version of the planning cycle was purposely chosen for this study. A more comprehensive version can be found in the *Guidelines for Education Sector Plan Preparation* (IIEP-UNESCO and GPE, 2015: 13).

Glossary

Regional assessments: Regional assessments 'are similar to international assessments, but they are administered in a given geographical world region. For example, SEACMEQ assessments are conducted in southern and eastern Africa' (IIEP-UNESCO, 2019a: 79).

Regression analysis: 'Regression analysis is a set of statistical methods used for the estimation of relationships between a dependent variable and one or more independent variables. It can be utilized to assess the strength of the relationship between variables and for modelling the future relationship between them' (Corporate Finance Institute, n.d.).

Representative sampling: 'Representative sampling is a type of statistical sampling that allows us to use data from a sample to make conclusions that are representative for the population from which the sample is taken' (D'Exelle, 2014).

Results-based management: 'A management strategy focusing on performance and achievement of outputs, outcomes and impacts' (OECD, 2002: 34).

Results chain: 'The causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts, and feedback' (OECD, 2002: 33).

Results framework: 'The program logic that explains how the development objective is to be achieved, including causal relationships and underlying assumptions.' (OECD, 2002: 33).

Sample-based assessment: An assessment conducted on a sample, i.e. 'a subset of units in a population, selected to represent all units in a population of interest. Testing a sample of a population is an effective and efficient way to gather information to describe performance across the

Glossary

education system (or a clearly defined part of it)' (GPE and ACER, 2019: 61).

Sampling: 'Scientific sampling methods ensure that the sample is reflective of the population and inferences about the population can be made from observations of the sample. This means that statements about the population can be based on the findings of the study conducted using the sample' (ACER-GEM; UIS, 2017: 28).

School grant: A school grant is the transfer of funds from the central level directly to schools to cover their running costs and to provide them with more autonomy in how their finances are managed. This is not the payment of school fees/tuition (IIEP-UNESCO and GPE, 2018).

Secondary education: 'Secondary education provides learning and educational activities, building on primary education and preparing for labour market entry, post-secondary non-tertiary education and tertiary education. Broadly speaking, secondary education aims at learning at an intermediate level of complexity. ISCED distinguishes between lower and upper secondary education (ISCED levels 2 and 3)' (UIS, 2012: 83).

Stakeholders: 'Agencies, organisations, groups or individuals who have a direct or indirect interest in the development intervention or its evaluation' (OECD, 2002: 35).

Standardized assessment: 'A test procedure in which the questions, papers, administration conditions, scoring and interpretation of results are applied in a consistent and pre-determined manner for all test-takers. Interpretation of standardized test scores can be norm-referenced or criterion-referenced. Norm-referenced standardised tests allow for comparisons of results between students which cannot reliably be inferred from non-standardised tests. Criterion-

Glossary

referenced standardised tests allow for identification of students who have attained a cut-score with respect to the skill or curriculum area being tested (the criterion), irrespective of the performance of their peers. Standardised scores allow for placement of students on a readily understandable scale, commonly centred at 100 to represent the average nationally standardised score for the population concerned. Such comparisons cannot meaningfully be made using raw scores or percentage scores' (DFID, 2011: 29).

Teachers' professional development: 'Professional development is defined as activities that develop an individual's skills, knowledge, expertise and other characteristics as a teacher' (OECD, 2009: 49).

Variance: 'The variance is the mean square deviation of the variable around the average value. It reflects the dispersion of the empirical values around its mean' (OECD, n.d.).

Annex 2. Methodology brief

This annex provides additional information on the methodology that framed the IIEP research programme on the 'Use of Learning Assessment Data in the Planning Cycle'. It first outlines the overall study methodology and then details the organization of data collection for each country that participated in the programme.

Overall research methodology

The project adopted a qualitative research approach to examining how different processes lead to the observed use of learning assessment data. A number of research methods were used.

Research methods

Interviews

In-depth, semi-structured interviews were conducted with selected respondents. It was important to gather different actors' perspectives, which vary, among other things, due to the diversity of their positions (e.g. planning specialists, assessment officers, curriculum developers) and their level of intervention (i.e. national, regional, district, school). The IIEP project team identified the following categories of interviewees:

— MoE officials and retired officers from different departments in charge of planning at central and decentralized levels;
— MoE staff/agencies/departments in charge of national examinations and national assessments, as well as

members of the international/regional assessment teams at central and decentralized levels;

- representatives of international partners involved in the funding/management of international/regional/national learning assessments;
- representatives of the technical agencies involved in LSAs;
- civil society representatives (e.g. NGOs and organizations involved in citizen-led assessments, teacher unions).

Interview schedules aimed to guide the discussion while allowing interviewees to address any topic relevant to the study. They were structured around three main themes:

- stakeholders' understanding of the intended use of learning data as defined in a regulatory framework and their knowledge of the existing regulatory documents;
- their knowledge of actual practices linked to the use of learning assessment data in the planning cycle;
- their insights on factors that determine the way these data are used, focusing on political economy and institutional analysis elements.

While these three parts structured all the interviews, it was important to ensure they were slightly adapted to the respondents' profiles (e.g. statistical elements were discussed in depth with teams in charge of the learning assessment data analysis; international partners were asked about their organizations' position on assessments).

The presence of two interviewers was required. While the first interviewer was supposed to lead the interview by asking questions, the second was expected to take notes

and provide support to the main interviewer when necessary (e.g. when clarification was needed, or a question was overlooked). However, not all interviews met this requirement for logistical reasons.

For confidentiality purposes, participants were only identifiable by a unique code. Interviewers recorded discussions on respondents' consent and then transcribed them.

Focus group discussions

Focus group discussions sought to gather information similar to that of interviews but generated through the interactions between the same group members. They were structured as per discussion guidelines. Questions were customized by actor group and by country, aiming to address issues that needed more information and that were not fully addressed during the individual interviews.

The research design initially targeted three categories of stakeholders:

— planning unit team;
— national assessment team;
— teacher representatives.

Ultimately, focus group discussions were not organized with teacher representatives, due to logistical reasons.

Focus group discussions took place at both central and decentralized administrative levels. They were organized with the same participants who had been interviewed individually. At central level, actors from different units

were sometimes invited to participate in the same focus group discussion to encourage dialogue between them on the topic and learn from their interactions.

Observations

The aim of the observations was to collect information on the interplay between actors. Observation of different events allowed researchers to explore dynamics that would not otherwise have been communicated by stakeholders in interviews or focus group discussions. It sought to understand whether learning data were valued and who advocated for their use. It also helped identify who was bringing learning data into the discussion and for what purposes. Observations focused on the use of learning data but general interactions between different actors were key elements for studying their dynamics. They took place during various events and meetings, such as ESP preparation workshops, local education group meetings, joint sector review meetings, etc.

An observation grid served as a guide for the analysis of the observation. It was structured according to the following sections:

- The event: the way actors intervene in the discussion, the weight of their participation, the themes discussed and mention of learning assessment data.
- Learning assessment data and evidence: description of the way data are brought into the discussion, discussed, and welcomed, the argument they are supposed to support.
- After observation: general observations after the event on actors' attitudes and discourse throughout the discussion.

Methodology brief

Documentary analysis

A desk review complemented other sources of information. Types of documents that were analysed included:

- education sector plans;
- education sector analyses;
- budget reports or evaluations;
- monitoring and evaluation documents;
- national and international studies;
- education laws;
- relevant policy documents/regulations;
- external programming documents;
- other relevant documents (e.g. PowerPoint presentations for specific events).

Data analysis

NVivo software used for data analysis facilitated the triangulation of information that had come from different sources. Research team members created nodes that corresponded to the sub-questions or sub-themes of the analytical framework, e.g. dissemination of assessment results. They then proceeded with data coding, i.e. they identified links between multiple data sources by assigning all information related to a specific theme to a corresponding node. Based on NVivo node extraction (exporting all information related to each specific sub-question or sub-theme), the research team prepared preliminary analyses for each country as per the same analytical framework that structured interview schedules.

Methodology brief

Detailed research methodology by country

1. The Gambia

Context of data collection

The data collection took place at a convenient time in The Gambia: an assessment policy had recently been developed that introduced changes to the assessment system (e.g. creation of the AU). It was also an opportunity to make updates to an important earlier study (Senghor, 2014) on the use of assessment data in The Gambia.

Organization of data collection

Two IIEP missions were organized to The Gambia (December 2018 and October 2019) and they respectively marked the beginning and the end of data collection in the country. Four officers from the MoBSE AU participated in data collection. Their involvement in the study was particularly useful, as they could analyse the uptake of the data they produced. However, their closeness to the learning assessment system might have introduced a certain bias, since they knew most of the study participants. Nevertheless, the benefits of their participation were considered to be higher than the potential risks.

At central level, the AU and a member of the IIEP research team conducted interviews, focus group discussions, and observation, while the AU handled interviews at decentralized level. All six Gambian regions were covered (i.e. Regions 1 through 6). Overall, the research team conducted 25 interviews at central level and nine at decentralized level, three focus group discussions, as well as one observation.

Methodology brief

Table A1. List of interviewees in The Gambia

ORGANIZATION	POSITION
Central level	
MoBSE	Monitoring & Evaluation – Principal Education Officer Director of Science and Technology Director of Basic and Secondary Education Administration & Finance – Principal Education Officer Donor Coordinator of the Project Coordination Unit (PCU) Director of Planning System Analyst – Principal Education Office Acting Project Manager Director of Human Resources
MoBSE CREDD	Director
MoBSE SQAD	Director Principal Education Officer
Permanent Secretary Office	Principal Education Officer
General Secretariat for Islamic and Arabic Education (GSIAE) or Amaanah	Secretary General
MRC Holland	Scholarship Coordinator
UNICEF	Education Programme Specialist
World Bank	Project Manager Consultant

Methodology brief

ORGANIZATION	POSITION
Effective Intervention	Research Manager
Education for All Campaign Network (EFANet)	National Coordinator
Forum for African Women Educationalists (FAWEGAM)	National Coordinator
Gambia Association of Teachers of English (GATE)	President
Gambia College	Head of School of Education
University of The Gambia	University Lecturer
Decentralized level	
MoBSE Regional Directorates	Director of Region 1 Director of Region 2 Director of Region 3 Principal Education Officer of Regional Education Directorate 3 Director of Region 4 Principal Education Officer of Regional Education Directorate 4 Principal Education Officer of Regional Education Directorate 5 Planning Officer of Regional Education Directorate 5 Director of Region 6 Acting Principal Education Officer Region 6

Methodology brief

Table A2. Focus group discussion participants in The Gambia

Focus group discussions
1. All four members of the AU (conducted by an IIEP researcher)
2. SQAD Director, SQAD Education Officer, Monitoring & Evaluation Officer, CREDD Principal Education Officer
3. Director of the Science and Technology Education Directorate (STED), Basic and Secondary Education Programme Director, Principal Education Officer of Region 1, Early Childhood Development Unit Principal Education Officer, INSET Education Officer

Table A3. Observation details in The Gambia

Meeting observed	Date and place	Objectives of the meeting	Participants
Local Education Group (LEG) Meeting	Regional Education Directorate 1, 8 October 2019	Regular LEG meeting to discuss pending education sector issues among partners and MoBSE officials	— UNICEF education specialist and deputy representative — Peace Corps — Post-Secondary Education coordinator — *Agence Belge de Développement* (ENABEL) — Action Aid — Monitoring and Evaluation director — EFANet — PCU manager — Deputy Permanent Secretary — MRC Holland — CREDD — Initiative for Strategic Litigation in Africa (ISLA) — MoBSE — Amanah — Anglican Diocese of Gambia — Catholic Education Secretariat (CES)

Methodology brief

Documents collected

Types of documents and sources used in The Gambia desk review include:

- education sector plans;
- education sector analysis;
- M&E documents;
- evaluation reports (e.g. EFA report, READ report) and EMIS reports (i.e. education statistics);
- assessment reports;
- education policy documents (e.g. assessment policy, curriculum framework, school management manual);
- school report cards;
- documents related to an institutional framework such as the assessment team job description, terms of reference for cluster monitors;
- MoBSE website;
- external programming documents (development partners' aides-memoires, external project evaluation reports, and project appraisal documents);
- international studies;
- presentations (e.g. World Bank's presentation on the uses of assessment information to support student learning in The Gambia).

2. Ghana

Context of data collection

The data collection at central level was organized together with a USAID consultant and with the logistical assistance of the USAID Ghana mission. USAID partnered with IIEP in two project countries (Ghana and Senegal) and both institutions shared data that were collected. Moreover, at the time of data collection, a new ESP (2018–2030) had been recently adopted in the country.

Organization of data collection

One joint IIEP and USAID mission was organized to Ghana in June 2019 to conduct interviews at central level. After this first mission, the IIEP organized the remaining data collection in Ghana with independent education consultants from MKA Education Services. Data collection spanned to November 2019. MKA Education Services conducted interviews and focus group discussions at decentralized level, as well as one observation at central level. Overall, 19 interviews at central level and 28 interviews at decentralized level, eight focus group discussions, and one observation were conducted in Ghana.

Consultants visited four regions and one district in each of these regions: Ashanti region and Ejisu district, Upper West region and Wa West district, Greater Accra region and La Nkwantanang Madina district, Volta region and Adaklu district.

Methodology brief

Table A4. List of interviewees in Ghana

ORGANIZATION	POSITION
Central level	
Ghana Education Service	Former Acting Director General
Ministry of Education	Assistant to the Minister of Education
Monitoring & Evaluation Unit	Principal Planning Officer
NaCCA	Retired official – Former Executive Secretary
NEAU	Acting Director
National Inspectorate Board (NIB)	Retired Executive Secretary
Statistics, Research, Information and Monitoring (SRIM)	Director
FHI360	Chief of Party, Learning Activity
IDP Foundation, Inc	Country Director
Innovations for Poverty Action (IPA)	Senior Education and Implementation manager
Perkins International	Regional Coordinator (Literacy for Visually Impaired Students)
UNESCO	National Education Programme Officer
UNICEF	Chief of Education
USAID	Senior Education Specialist
World Bank	Senior Education Consultant
Centre for Democratic Development (CDD)	Team Leader for Social Accountability and SDGs programming and advocacy
Ghana National Education Campaign Coalition (GNECC)	National Coordinator
Northern Network for Education Development (NNED)	Coordinator

Methodology brief

ORGANIZATION	POSITION
Decentralized level	
Adaklu District Educational Office	District Inspection Officer District Exam Officer District Statistics Officer District ICT/EMIS officer
Ashanti Regional Education Office	Regional Planning Officer Regional EMIS Officer Regional Statistics Officer Regional Inspectorate Officer
Eastern Regional Directorate	Director
Ejisu Municipal Education Office	District Planning Officer District EMIS Officer District Statistics Officer District Assessment Officer
Greater Accra Regional Education Office	Regional EMIS Officer Regional Training Officer Regional Inspectorate Officer
La Nkwantana Madina Municipal Education Office	Municipal Exam Officer Municipal Planning Officer Municipal Research/Statistics Officer Municipal Training Officer
Upper West Regional Education Office	Regional Assessment Officer Regional Statistics/EMIS Officer Regional Inspectorate Officer

Methodology brief

ORGANIZATION	POSITION
Volta Regional Education Office	Regional Planning Officer Regional Statistics Officer Regional Basic Education Officer
Wa West District Education Office	District Planning Officer District EMIS Officer Regional Inspectorate Officer

Table A5. Focus group discussion participants in Ghana[31]

Focus group discussions sites
1. Greater Accra Regional Education Office
2. La Nkwantana Madina Municipal Education Office
3. Volta Regional Education Office
4. Adaklu District Educational Office
5. Ashanti Regional Education Office
6. Ejisu Municipal Education Office
7. Upper West Regional Education Office
8. Wa West District Education Office

31 Focus group discussions took place at eight sites, where they gathered respondents from municipal, district, or regional education offices who had been interviewed individually beforehand. Thus, as listed above, in each of the sites, focus group discussions were conducted with officers in charge of: the planning unit, the EMIS unit, the inspectorate unit, the research/statistics unit, and the examination and training unit at both the Regional and Municipal/District Education Offices. It must be noted that officers who had been first interviewed individually were not always present at the subsequent focus group discussions, hence officers from other units who were also part of the planning committee at the region/municipal/district offices represented them.

Table A6. Observation details in Ghana

Meeting	Date and place	Objectives of the meeting	Participants
Education Development Partners' Monthly Meeting	UNESCO office, 19 February 2020	Inform participants about the latest developments in the education sector in Ghana and help put the activities of different development partners into perspective	— Department for International Development (DFID) Ghana — Embassy of France — MasterCard Foundation — Japan International Cooperation Agency (JICA) — UNESCO — UNICEF — USAID — Ministry of Education — GES — Planning, Budgeting, Monitoring and Evaluation Division (PBME) — UNHCR

Documents collected

Types of documents and sources used in the Ghana desk review include:

— education sector plans;
— M&E documents;
— education sector analyses;
— assessment reports;
— external project evaluations, external country reports, and project information documents;
— international studies;
— presentations (e.g. prepared for the national education week).

3. Guinea

Context of data collection

In Guinea, the IIEP study was conducted at the same time as the country was developing its education strategy for the upcoming ten years, Programme Décennal de l'Éducation en Guinée (ProDEG) 2020–2029. Research missions took place in parallel with the IIEP Technical Cooperation Unit's missions to support the ProDEG development, thus creating an opportunity to observe how the programme was progressing.

In addition, at the time of data collection, Guinea was preparing for the PASEC 2019 administration.

Organization of data collection

Two missions were organized to Guinea (November 2018 and February 2019) and data collection took place from November 2018 to June 2019. The IIEP research team member conducted interviews and observations at central level with logistical support from a ministry of education official. The Head of Advanced Studies of the Guinean Institute for Education Sciences (ISSEG) conducted interviews and focus group discussions at decentralized level.

Four regions were visited for data collection: Boké, Faranah, Kindia and Labé.

Overall, 20 interviews at central and 16 interviews at decentralized level, four focus group discussions, and three observations were conducted in Guinea.

Methodology brief

Table A7. List of interviewees in Guinea

ORGANIZATION	POSITION
Central level	
Directorate of Secondary, General and Vocational Education (DESGT)	Director
General Inspectorate of Education (IGE)	General Inspector
Medium-Term Expenditure Framework (CDMT)	Director Associate Officer
Ministry of Higher Education and Scientific Research (MESRS)	Cabinet Officer
National Directorate for Literacy, Non-Formal Education and Promotion of National Languages (DNAPL)	Director
National Directorate of Basic Education	Director
National Institute for Pedagogical Research and Action (INRAP)	Deputy Director EGRA Programme Officer
National Service for School Examination and Course Guidance (SNESCO)	Deputy Director and Human Resources Manager
National Service for the Evaluation of the Education System (SNESE)	Director Specialist
National Unit for ESP Monitoring (CN/PSE)	Procurement Officer Head of Unit
Strategic Expenditure Office of the Ministry of Education and Literacy (MENA/BSD)	Director Head of the 'Studies and Planning' unit Technical and IT Assistant

Methodology brief

ORGANIZATION	POSITION
Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ Guinée)	Senior Cooperation and Partnership Expert – Basic Education Programme
UNICEF	Chief of Education Education Specialist
Decentralized level	
Prefectural Direction of Education Boké	Head of the Strategic Expenditure Office
Regional Inspectorate of Education Boké	Planning Officer Deputy General Inspector Regional Inspector
Prefectural Direction of Education (DPE) Faranah	Chief of BSD
Regional Inspectorate of Education Faranah	Chief of BSD Regional Inspector
Prefectural Direction of Education (DPE) Kindia	Chief of Unit BSD
Regional Inspectorate of Education Kindia	Regional Inspector Chief of BSD/Statistician Officer Head of Pedagogical Section, Secondary, General and Technical Education / Deputy Inspector Chief of Unit BSD
Prefectural Direction of Education (DPE) Labé	Chief of Unit BSD
Regional Inspectorate of Education Labé	Chief of Unit BSD Head of Pedagogical Section, Secondary, General and Technical Education Regional Inspector

Methodology brief

Table A8. Focus group discussion participants in Guinea

Focus group discussions: Units
1. Regional Inspectorate of Education Boké
2. Regional Inspectorate of Education Faranah
3. Regional Inspectorate of Education Kindia
4. Regional Inspectorate of Education Labé

Methodology brief

Table A9. Observation details in Guinea

Meeting	Date and place	Objectives of the meeting	Participants
Workshop for the preparation of ProDEG	Ministry of Education and Literacy (MENA), 29 November 2018	Present the simulation model to leaders of the three ministries of education for discussion and approval of the main action lines	— UNICEF — GPE — GIZ — French Development Agency (AFD) — Permanent Secretaries of the three ministries of education — IIEP-UNESCO — National technical team — Representatives of MENA
Workshop for the preparation of the programmes and sub-programmes of the ten-year Education Programme of Guinea (ProDEG)	Kindia, 4–7 February 2019	Prepare the programmes and sub-programmes of the ProDEG	— Representatives from the three ministries of education and their departments — SNCESE — UNICEF — AFD — IIEP-UNESCO
Workshop for validation of the sector analysis and the state of progress in ProDEG development	Conakry, 8 February 2019	Validate the sector analysis and state of progress of ProDEG	— AFD — World Bank — GIZ — IIEP-UNESCO — UNICEF — National technical team who participated in the preparation of ProDEG

Documents collected

Types of documents and sources used in the Guinea desk review include:

- country development plan;
- education sector plans;
- M&E documents;
- education sector analyses;
- assessment reports;
- regulatory texts;
- joint sector reviews aides-memoires;
- external project evaluations, external country reports, and project appraisal documents;
- international studies.

4. Namibia

Organization of data collection

Data collection took place from March 2019 to July 2019 and one IIEP mission was organized (in March 2019). The IIEP research team member was assisted by two researchers from the University of Namibia to conduct interviews at central level. National researchers then conducted the rest of data collection at central (remaining interviews, focus group discussions, observations) and decentralized levels (i.e. regions). They collected data in seven out of 14 Namibian regions: Khomas, Kavango, Hardap, Erongo, Omusati, Oshana, Zambezi.

Methodology brief

Overall, the team conducted: 13 interviews at central level and 19 interviews at decentralized level, two focus group discussions, and three observations.

Table A10. List of interviewees in Namibia

Organization	Position
Central level	
Directorate of Educational Planning and Development (PAD)	Director of Planning and Development Deputy Director of EMIS division Senior Education Officer in EMIS division
DNEA	Deputy Director of Data and Certification Deputy Director of Examinations, Research &Development Chief Education Officer of Research and Development Sub-Division Senior Education Officer of Research and Development Sub-Division
NIED	Chief Education Officer of Curriculum Division
Programmes and Quality Assurance (PQA)	Chief Education Officer Deputy Director
UNICEF	Integrated Education Officer
UNESCO	Programme Specialist
EU	Representative

Methodology brief

Organization	Position
Decentralized level	
Erongo region	REXO Chief Education Planner Acting CEO Professional Development
Hardap region	Chief Education Planner Chief Education Officer, Professional Development
Kavango region	Chief Education Planner Chief Education Officer, Professional Development Regional Examination Officer
Khomas region	REXO Chief Education Officer, Professional Development Chief Education Planner
Omusati region	Chief Education Planner Acting Regional Examination Officer Chief Education Officer, Professional Development
Oshana region	Chief Education Officer, Professional Development Regional Examination Officer Chief Education Planner
Zambezi region	Chief Education Planner REXO

Methodology brief

Table A11. Focus group discussion participants in Namibia

Focus group discussions: Units involved
1. Assessment team (DNEA)
2. Planning Unit (PAD)

Observations

Table A12. Observation details in Namibia

Meeting	Date and place	Objectives for the meeting	Participants
Annual Review Meeting	Ministry of Education, arts, and Culture (decentralized level), 2 April 2019	Review the progress made in implementation of the 2018/19 annual plans for three decentralized directorates of education (Hardap, Oshana, Omusati regions)	Executive Director, Deputy Executive Director, Directors of Education, Deputy Directors and Chief Planners
Validation of 2019/20 Annual Plans of the National Directorates of Education	Ministry of Education, Arts, and Culture (central level), 3 April 2019	Validate the draft 2019/20 annual plans of the directorates of education to be in line with the ministry of education, arts and culture Strategic Plan 2018–2022	Executive Director, Deputy Executive Director, Directors of Education, Deputy Directors, and Chief Planners

Meeting	Date and place	Objectives for the meeting	Participants
Validation of 2019/20 Annual Plans of the Decentralized Directorates of Education	Ministry of Education, Arts, and Culture (central level), 4 April 2019	Validate 2019/20 Annual Plans of the decentralized directorates of education	Executive Director, Deputy Executive Director, Directors of Education, Deputy Directors, and Chief Planners

Documents collected

Types of documents and sources used in the Namibia desk review include:

- country development plan;
- education sector plans;
- M&E documents;
- education sector analyses;
- education policy documents;
- teachers' professional development reports;
- assessment reports;
- regulatory texts;
- international studies.

5. Senegal

Context of data collection

Data collection at central level was organized together with a USAID consultant and with logistical help from the USAID Senegal mission. USAID partnered with IIEP in two project countries (Ghana and Senegal) and both institutions shared data that were collected.

Organization of data collection

The data collection took place from June 2019 to March 2020. One joint IIEP and USAID mission was organized to Senegal in June 2019 to conduct interviews at central level. After this first mission, the IIEP organized the remaining data collection with an independent education consultant. A second IIEP mission to Senegal took place in January 2020 to conduct the remaining interviews at central level, as well as focus group discussions together with a consultant. She then carried out interviews at decentralized level.

Overall, 25 interviews at central level, 21 interviews at decentralized level, two focus group discussions, and one observation were conducted in Senegal.

Methodology brief

Table A13. List of interviewees in Senegal

Organization	Position
Central level	
Ministry of Education (MENA)	Secretary General
General Inspectorate of Education	General Inspector
Directorate of Educational Planning and Reform (DPRE/MoE)	Director
Directorate of Primary Education (DEE)	Director
Directorate of General Secondary Education (DEMSG)	Coordinator
Directorate of Examinations and Competitions (DEXCO)	Director
IT and Management System of Education (SIMEN)	Coordinator
National Research Institute for Education (INEADE)	Director Retired Official New Director
Directorate for Reform and Education Planning (DPRE)	Retired Official (National Observatory on Parity)
Programme for Basic Education Improvement in Casamance (MoE PAECBA)	Retired Official/PAECBA Coordinator
Directorate for Early Childhood Education (DEPS)	Director
Directorate for Early Childhood Education (DEPS)	Coordinator
PASEC	Coordinator
Chemonics (*Lecture pour Tous*)	Senior Monitoring, Evaluation and Learning Specialist

Methodology brief

Organization	Position
USAID	Education Evaluation Specialist
IIEP Dakar	Quality Management Programme Coordinator Thematic Group on Education Teaching and Learning Researcher
World Bank	Senior Education Specialist
Coalition of Organizations in Synergy for the Defense of Public Education (COSYDEP)	Coordinator
Research Laboratory on Economic and Social Transformations (LARTES-IFAN)	Coordinator
University Cheikh Anta Diop	Monitoring and Evaluation Specialist
Decentralized level	
Academy Inspectorate Dakar	Head of the office for pedagogical management and learning assessment Planning officer Head of Statistics, Monitoring & Evaluation Office Inspector
Academy Inspectorate Kaolack	Secretary General Monitoring & Evaluation manager Planning officer Administrator/Former planning manager Inspector

Methodology brief

Organization	Position
Academy Inspectorate Thiès	Inspector Head of the office for pedagogical management and learning assessment Former Secretary General Planning officer
Inspectorate for Education and Training (IEF) Diourbel	Inspector Planning officer Assistant/Monitoring & Evaluation manager Secretary General
Inspectorate for Education and Training (IEF) Fatick	Planning officer Statistics officer Monitoring & Evaluation manager Inspector
Inspectorate for Education and Training (IEF) Kaolack-Commune	Inspector

Table A14. Focus group discussion participants in Senegal

Focus Group Discussions: Units
1. Assessment team (INEADE)
2. Planning Unit (DPRE)

Methodology brief

Table A15. Observation details in Senegal

Meeting	Date and place	Objectives of the meeting	Participants
Meeting on the IIEP-UNESCO Dakar's support within the framework of the implementation of PADES[32] project activities	MoE DPRE, February 2020	Elaborate the dashboard and vulnerability mapping and prepare the review	— DPRE — SIMEN — PADES — IIEP Dakar — AFD

Documents collected

Types of documents and sources used in the Senegal desk review include:

— country development plan;
— policy statement;
— education sector plans;
— M&E documents;
— education sector analyses;
— education policy and other national programming documents (Palme project);
— teachers' professional development reports;
— assessment reports;
— regulatory texts;
— performance contracts;
— joint sector review aide-memoire;

32 The support programme for the development of education in Senegal (*Programme d'Appui au Développement de L'Éducation au Sénégal*).

- project evaluation report, project appraisal documents, and external country reports;
- school management questionnaires (UNICEF-conducted);
- national (i.e. country statistics report) and international studies;
- presentations (e.g. a review of the status of evaluations in the PAQEEB framework).

6. Zambia

Organization of data collection

The data collection took place from January 2020 to March 2020. It was carried out entirely by two researchers from the University of Zambia at central and decentralized levels. Researchers collected data in four provinces (out of ten).

Overall, 20 interviews at central level, 12 interviews at decentralized level, three focus group discussions, as well as one observation were conducted in Zambia.

Methodology brief

Table A16. List of interviewees in Zambia

Organization	Position
Central level	
Ministry of General Education	Assistant Director for Research and Statistics Assistant Director Planning/Chief Planner responsible for the budget and planning Director Standards Interim Project Coordinator Planning Officer responsible for M&E implementation plan Principal Planner in charge of budget and projects Retired Chief Education Officer for Teacher Education Statistician
ECZ	Director
Ministry of Planning	Principal Planning Officer
Teaching Council of Zambia	Registrar Teaching Council
Devtech	Project Manager
UNICEF	Chief of Education
USAID	Deputy Chief of Party of Reading Project & Retired Chief Planning Officer at the Ministry of Planning Education Director
World Bank	ZEEP Project coordinator
Zambia National Commission for UNESCO	Programme Specialist

Methodology brief

Organization	Position
Zambia National Education Coalition (ZANEC)	Chief Schools Officer – Executive Director Executive Director
Decentralized level	
Ministry of General Education Central Province	Chief Planning Officer Education Officer Teacher Education Principal Education Standard Officer
Ministry of General Education Copperbelt Province	Education Officer Teacher Education Senior Education Standard Officer Planning Officer
Ministry of General Education Lusaka region	Teacher Education Officer Provincial Resource Centre Coordinator
Ministry of General Education Southern Province	Education Officer (In-service Training and Teacher Training coordination) Senior Education Standard Officer Mathematics Senior Education Standard Officer Language
Ministry of General Education Southern Province	Senior Planner

Table A17. Focus group discussion participants in Zambia

Focus Group Discussion Sites
1. Principal Education Office in Choma (Southern Province)
2. Principal Education Office in Kabwe (Central Province)
3. Principal Education Office in Ndola (Copperbelt Province)

Table A18. Observation details in Zambia

Meeting	Date and place	Objectives of the meeting	Participants
EU data validation workshop	N.A.	Discuss findings of the sub-sector analysis of Zambian education sector	— Ministry of Education representatives — Teaching Council — European Union delegates — UNESCO — UNICEF — World Bank — DFID — USAID — British Council — Oxfam

Documents collected

Types of documents and sources used in the Zambia desk review include:

— country development plan;
— policy statement;
— education sector plans;
— M&E documents;
— education sector analyses;
— implementation documents (e.g. ESP implementation framework);
— education policy (e.g. NLAF) and other national programming documents;
— teachers' professional development reports;
— assessment reports;
— terms of references of the data management committee;

Methodology brief

- external programming documents and external country reports;
- national (e.g. literature review of best practices for improving learner performance in Zambian schools) and international studies;
- presentations (e.g. a presentation on student assessment in Zambia at a World Bank symposium).

References

ACER-GEM (Australian Council for Educational Research, Centre for Global Education Monitoring); UIS (UNESCO Institute for Statistics). 2017. Principles of Good Practice in Learning Assessment. The ACER Centre for Global Education Monitoring. http://uis.unesco.org/sites/default/files/documents/principles-good-practice-learning-assessments-2017-en.pdf.

ADB (Asian Development Bank). 2015. *Third Primary Education Development Project Additional Financing: Report and recommendation of the president*. https://www.adb.org/projects/documents/third-primary-education-development-project-additional-financing-rrp.

Addey, C.; Gorur, R. 2020. 'Translating PISA, Translating the World'. In: *Comparative Education,* 56(4), 547–564. https://doi.org/10.1080/03050068.2020.1771873.

Addey, C.; Sellar, S. 2019. 'Is it Worth it? Rationales for (Non)Participation in International Large-Scale Learning Assessments'. *Education Research and Foresight Working Papers Series*, 24. Paris, UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000368421.

Addey, C.; Sellar, S.; Steiner-Khamsi, G.; Lingard, B.; Verger, A. 2017. 'The Rise of International Large-Scale Assessments and Rationales for Participation'. In: *Compare: A Journal of Comparative and International Education,* 47(3), 434–452. https://doi.org/10.1080/03057925.2017.1301399.

Auld, E.; Rappleye, J.; Morris, P. 2019. 'PISA for Development: How the OECD and World Bank shaped education governance post-2015'. In: *Comparative Education.* 55(2), 197–219. https://doi.org/10.1080/03050068.2018.1538635.

References

Benveniste, L. 2002. 'The Political Structuration of Assessment: Negotiating state power and legitimacy'. In: *Comparative Education Review,* 46(1), 89–118. https://doi.org/10.1086/324051.

Breakspear, S. 2014. *How Does PISA Shape Education Policy Making? Why how we measure learning determines what counts in education.* Victoria: Centre for Strategic Education. Available at: https://allchildrenlearning.org/wp-content/uploads/2019/11/Breakspear-PISA-Paper.pdf.

Britannica. 2016. The Editors of Encyclopaedia. "inference". *Encyclopedia Britannica,* 26 Oct. www.britannica.com/science/inference-statistics.

Bryman, A. 2004. *Social Research Methods*. Second Edition. Oxford: Oxford University Press.

Clarke, M. 2012. *What Matters Most for Student Assessment Systems: A framework paper*. Systems Approach for Better Education Results (SABER) student assessment working paper, 1. Washington DC: World Bank. https://openknowledge.worldbank.org/handle/10986/17471.

Coombs, P.H. 1970. *What Is Educational Planning?* Fundamentals of Educational Planning, 1. Paris: IIEP-UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000076671/PDF/076671engo.pdf.multi.

Corporate Finance Institute. n.d. 'Regression Analysis: Formulas, explanation, examples and definitions'. https://corporatefinanceinstitute.com/resources/knowledge/finance/regression-analysis/.

D'Exelle, B. 2014. 'Representative Sample'. In *Encyclopedia of Quality of Life and Well-Being Research*, edited by Alex C. Michalos, 5511–13. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-0753-5_2476.

References

DFID (Department for International Development, UK). 2011. 'National and International Assessments of Student Achievement'. Guidance Note: A DFID practice paper. UK Department for International Development. https://assets.publishing.service.gov.uk/media/57a08ab440f0b6497400071a/04-Learning-Outcomes-How-To-Note.pdf.

Dubeck, M.M.; Gove, A. 2015. 'The Early Grade Reading Assessment (EGRA): Its theoretical foundation, purpose, and limitations'. In: *International Journal of Educational Development,* 40, 315–322. https://doi.org/10.1016/j.ijedudev.2014.11.004.

Elks, P. 2016. The Impact of Assessment Results on Education Policy and Practice in East Africa. DFID Think Piece. Health and Education Advice and Resource Team (HEART).

Fischman, G.E.; Marcetti Topper, A.; Silova, I.; Goebel, J.; Holloway, J.L. 2019. 'Examining the Influence of International Large-Scale Assessments on National Education Policies'. In: *Journal of Education Policy,* 34 (4), 470–499. https://doi.org/10.1080/02680939.2018.1460493.

Furiv, U. Unpublished manuscript. '[Draft] Twenty Years of Discourse on the Use of Learning Assessment Systems in the Education Sector Plans (ESPs) of 46 Sub-Saharan African Countries'.

The Gambia. 2015. Ministry of Basic and Secondary Education. *Assessment Policy 2015–2022 (Second Draft)*.

— — — —. 2017. *Report on National Assessment Tests 2017 (Grades 3 and 8)*.

Gardner, E. 1989. 'Five Common Misuses of Tests'. ERIC Digest No. 108. http://ericae.net/db/digs/ED315429.htm.

GEM (Global Education Monitoring) Report Team. 2007. *Strong Foundations: Early childhood care and education. EFA*

References

Global Monitoring Report, 2007. Paris: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000147794.

———. 2008. *Education for All by 2015: Will we make it? EFA Global Monitoring Report, 2008*. Paris: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000154743.

———. 2017. *Accountability in Education: Meeting our commitments*. Global Education Monitoring Report, 2017/18. Paris: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000259338.

———. 2014. *Ghana 2013 Early Grade Reading Assessment and Early Grade Mathematics Assessment: Reports of findings*. https://earlygradereadingbarometer.org/pdf/Ghana_2013.pdf.

Ghana. 2014. Ministry of Education. Ghana Education Service. National Education Assessment Unit. *Ghana 2013 National Education Assessment: Summary of results*. https://shared.rti.org/content/ghana-2013-national-education-assessment-summary-results.

———. 2016. *Ghana 2016 National Education Assessment: Report of findings*. https://sapghana.com/data/documents/2016-NEA-Findings-Report_17Nov2016_Public-FINAL.pdf.

———. 2018. Ghana Ministry of Education. *Education Sector Analysis*. https://assets.globalpartnership.org/s3fs-public/2019-05-ghana-education-sector-analysis.pdf.

———. 2019. Ghana Ministry of Education. *Education Strategic Plan 2018–2030*. https://planipolis.iiep.unesco.org/sites/default/files/ressources/ghana-education-strategic-plan-2018-2030.pdf.

GPE (Global Partnership for Education). 2015. *The GPE Funding Model: A results-based approach for the education*

References

sector. Factsheet. Washington, DC: GPE. https://www.globalpartnership.org/sites/default/files/2015-06-gpe-funding-model_0.pdf.

———. 2017. *How GPE Supports Teaching and Learning*. Policy brief. Washington, DC: GPE. https://www.globalpartnership.org/sites/default/files/2017-09-gpe-policy-brief-teaching-learning.pdf.

———. 2019a. *Global Partnership for Education Results Report 2019*. Washington, DC: GPE. https://www.globalpartnership.org/sites/default/files/2019-06-10-gpe-results-report-2019.pdf.

———. 2019b. *Harnessing the Power of Data to Improve Education Systems*. https://assets.globalpartnership.org/s3fs-public/document/file/2019-02-GPE-factsheets-Data.pdf.

———. 2019c. *Strengthening Learning Assessment Systems: A knowledge and innovation exchange (KIX) Discussion Paper*. GPE. https://www.globalpartnership.org/sites/default/files/document/file/2019-07-kix-las-final.pdf.

———. 2020. *Country-Level Guide: Recommended education sector and GPE grants processes*. Washington, DC: GPE. https://www.globalpartnership.org/sites/default/files/document/file/2020-02-GPE-country-level-guide.pdf.

———. 2022. 'How we work'. GPE website: www.globalpartnership.org/what-we-do/how-we-work.

———. n.d. *English/French Glossary of the Global Partnership for Education (GPE)*. https://www.globalpartnership.org/content/gpe-glossary.

GPE; ACER (Australian Council for Educational Research). 2019. *Analysis of National Learning Assessment Systems [ANLAS]. Manual*. https://www.globalpartnership.org/sites/default/files/docs/2020-11-GPE-ANLAS-manual.pdf.

References

Guinea. 2016. Ministère de l'Enseignement Pré-Universitaire et de l'Alphabétisation. Service National de Coordination de l'Evaluation du Système Educatif. *Évaluation des Acquis des Élèves du CE2 en Français et Calcul : Rapport Technique*.

— — — —. 2014a. Ministères en charge du secteur de l'éducation. *Programme Sectoriel de l'Éducation 2015-2017*. https://planipolis.iiep.unesco.org/sites/default/files/ressources/education-sector-plan-guinea.pdf.

— — — —. 2014b. *Rapport Bilan du PSE 2008-2014*. Conakry.

— — — —. 2013. Ministère de l'Enseignement Pré-Universitaire et de l'Alphabétisation. Service National de Coordination de l'Evaluation du Système Educatif. *Evaluation des Compétences des Élèves de CE2 en Français et Calcul*.

— — — —. 2007. Ministère de l'Éducation Nationale et de la Recherche Scientifique. *Description du Programme Sectoriel de l'Éducation (PSE) 2008-2015.* Conakry. https://riseprogramme.org/sites/default/files/inline-files/Johnson_Assessing%20the%20assessments_RISE%20version_v2.pdf.

IIEP-UNESCO. 2019a. 'Research Guidelines. Use of Learning Assessment Data in the Planning Cycle'. Unpublished.

— — — —. 2019b. *Issue Brief 5: Using Data to Improve the Quality of Education*. IIEP-UNESCO Learning Portal. 5 May 2019. https://learningportal.iiep.unesco.org/en/issue-briefs/monitor-learning/using-data-to-improve-the-quality-of-education.

IIEP-UNESCO; GPE. 2015. *Guidelines for Education Sector Plan Preparation*. Paris: IIEP-UNESCO; Washington, D.C.: GPE. https://unesdoc.unesco.org/ark:/48223/pf0000233767.

— — — —. 2018. *Designing and Implementing a School Grant Policy*. https://unesdoc.unesco.org/ark:/48223/pf0000265168.

References

IIEP-UNESCO; World Bank; UNICEF. 2014. *Education Sector Analysis Methodological Guidelines. Volume 1. Sector-wide analysis, with emphasis on primary and secondary education.* Dakar: IIEP-UNESCO (Dakar office). https://unesdoc.unesco.org/ark:/48223/pf0000230532.

Irving, M.; Mitra, R. 2019. 'Strengthening Data Systems through Investing in Knowledge and Innovation'. 2019. Global Partnership for Education Blogs. https://www.globalpartnership.org/blog/strengthening-data-systems-through-investing-knowledge-and-innovation.

Johnson, D.; Parrado, A. 2020. 'Assessing the Assessments: Taking stock of learning outcomes data in India'. https://riseprogramme.org/sites/default/files/inline-files/Johnson_Assessing%20the%20assessments_RISE%20version_v2.pdf.

Kamens, D.H.; McNeely, C.L. 2010. 'Globalization and the Growth of International Educational Testing and National Assessment'. In: *Comparative Education Review,* 54 (1), 5–25. https://doi.org/10.1086/648471.

Kellaghan, T., Greaney, V. 2003. 'Monitoring Performance: Assessment and Examinations in Africa'. Paper commissioned for the ADEA Biennial Meeting 2003. https://biennale.adeanet.org/2003/papers/2D_Monitoring_ENG_final.pdf.

Kellaghan, T.; Greaney, V.; Murray, S. 2009. *Using the Results of a National Assessment of Educational Achievement*. National Assessments of Educational Achievement, v. 5. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/2667.

Kingdon, G.G.; Little, A.; Aslam, M.; Rawal, S.; Moe, T.; Patrinos, H.; Beteille T.; Banerji, R.; Parton, B.; Sharma, S.K. 2014. *A Rigorous Review of the Political Economy of Education*

References

Systems in Developing Countries. Final Report. Education Rigorous Literature Review. UK Department for International Development. https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/Political%20economy%202014Kingdon.pdf?ver=2014-04-24-141259-443

Lai, M. K.; Schildkamp, K. 2016. 'In-Service Teacher Professional Learning: Use of assessment in data-based decision-making'. In: G.T.L. Brown and L.R. Harris (eds), *Handbook of Human and Social Conditions in Assessment* (pp. 77–94). Routledge: New York.

Levine, V. 2013. *Education in Pacific Island States: Reflections on the failure of 'grand remedies'*. Pacific Islands Policy, 8. Honolulu: East-West Center. https://www.eastwestcenter.org/publications/education-in-pacific-island-states-reflections-the-failure-grand-remedies.

Lockheed, M. 2013. 'Causes and Consequences of International; Assessments in developing countries'. In: Meyer, H.-D. and Benavot, A. (eds), *PISA, Power, and Policy. The Emergence of Global Educational Governance* (pp. 163–79. Oxford: Symposium Books.

— — — —. 2016. *Measures That Matter: Learning Outcome Targets for Sustainable Development Goal 4: An examination of national, regional and international learning assessments*. Background paper prepared for the 2016 Global Education Monitor Report, Education for People and Planet: Creating Sustainable Futures for All. https://unesdoc.unesco.org/ark:/48223/pf0000245842.

'Multivariate Regression Analysis: Stata data analysis examples'. n.d. UCLA: Statistical Consulting Group. Last accessed 31 March 2021: https://stats.idre.ucla.edu/stata/dae/multivariate-regression-analysis/.

References

Murimba, S. 2005. 'The Impact of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)'. In: *Prospects*, 35 (1), 92–108. https://doi.org/10.1007/s11125-005-6819-7.

Namibia. 2015. Ministry of Education, Arts and Culture. *The SACMEQ IV Project in Namibia: A study of the conditions of schooling and the quality of primary education in Namibia*. SACMEQ Policy Research Report 4. Windhoek. www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iv/national-reports/final_sacmeq_iv_report_namibia-compressed-compressed.pdf.

———. 2017. Ministry of Education, Arts and Culture. *Situation Analysis: Basic education background to the strategic plan for the Basic Education Sector Plan 2017–2022*. Second Draft.

———. 2018. Ministry of Education, Arts and Culture. *Dissemination of Grade 5 Baseline Standardised Achievement Tests Reports*. Circular: DNEA 25/2018. Windhoek: Ministry of Education, Arts and Culture.

Nevo, I.; Slonim-Nevo, V. 2011. 'The Myth of Evidence-Based Practice: Towards evidence-informed practice'. In: *British Journal of Social Work,* 41 (6), 1176–1197. https://doi.org/10.1093/bjsw/bcq149.

OECD (Organisation for Economic Co-operation and Development). n.d. 'Glossary of Statistical Terms'. https://stats.oecd.org/glossary/detail.asp?ID=4337.

———. 2002. *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: OECD Publications. https://www.oecd.org/dac/2754804.pdf.

———. 2009. 'The Professional Development of Teachers'. In: *Creating Effective Teaching and Learning Environments:*

References

First results from TALIS, 47–86. Paris: OECD. https://www.oecd.org/berlin/43541636.pdf.

— — — —. 2015. *Pisa for Development: Capacity Needs Analysis: Zambia*. Paris: OECD. https://www.oecd.org/pisa/aboutpisa/NEW_Pisa%20for%20Development_Zambia_FINAL_revised.pdf.

— — — —. 2016. *PISA for Development, Capacity-Building Plan: Zambia*. Paris: OECD. https://www.oecd.org/pisa/aboutpisa/CBP%20Zambia_FINAL.pdf.

Paviot, L.; Saito, M. 2015. *From Assessment to Actions: Impact of student assessment data on educational policy reform*. [PowerPoint slides]. https://www.slideshare.net/IIEP_UNESCO/from-assessment-to-action.

Porter, C. 2010. *What Shapes the Influence Evidence Has on Policy? The role of politics in research utilisation*. Working Paper, 62. Oxford: Young Lives, Department of International Development, University of Oxford. https://www.younglives.org.uk/sites/default/files/migrated/YL-WP62-Porter-Evidence-and-Policy.pdf.

Postlethwaite, T.N.; Kellaghan, T. 2008. *National Assessments of Educational Achievement*. Education Policy Series, 9. Paris and Brussels: IIEP-UNESCO and IAE. https://unesdoc.unesco.org/ark:/48223/pf0000181753.

Ramírez, M.J. 2018. Quick Guide No. 2: *Making the Case for a Learning Assessment*. Montreal: UNESCO-UIS. https://unesdoc.unesco.org/ark:/48223/pf0000265404.

Raudonytė, I. 2019. *Use of Learning Assessment Data in Education Policy-Making*. IIEP Working Papers, April 2019. Paris: IIEP-UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000367608.

References

Raudonytė, I.; Bodin, M. 2021. *The Use of Learning Assessment Data: Zambia*. IIEP policy brief; the use of learning assessment data, 6. Paris: IIEP-UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000375498.

READ (Russian Education Aid For Development). 2010. 'Zambia Country Self Diagnosis'. Unpublished internal report.

Rizvi, F.; Lingard, B. 2010. *Globalizing Education Policy*. London; New York, NY: Routledge.

Robertson, S. 2018. *Reporting and Disseminating Large-Scale Assessment Results to a Variety of Stakeholders (Pt. 3)*. Presented at the TALENT workshop, Dakar, November 27.

Saito, Mioko. 1999. 'A generalizable Model for Educational Policy Research in Developing Countries'. Paris: IIEP-UNESCO. Available at: https://cice.hiroshima-u.ac.jp/wp-content/uploads/publications/Journal2-2/2-2-9.pdf.

Saito, M.; van Cappelle, F. 2010. 'Monitoring the Quality of Education: Exploration of concept, methodology, and the link between research and policy'. In: Wiseman, A.W. (ed), The Impact of International Achievement Studies in National Educational Policymaking (*International Perspectives on Education and Society*, Vol. 13 (pp. 3–34). Bingley: Emerald Group Publishing Limited. https://doi.org/10.1108/S1479-3679(2010)0000013004.

Schwantner, U.; de Chaisemartin, T. 2017. 'Ensuring Learning Data Matters'. Education for All blog, 9 February 2017. www.globalpartnership.org/blog/ensuring-learning-data-matters.

Senegal. 2013. Ministère de la Femme, de l'Enfant et de l'Entrepreneuriat féminin; Ministère de l'Éducation Nationale; Ministère de l'Enseignement Supérieur et de la Recherche; Ministère de la Formation professionnelle, de l'Apprentissage et de l'Artisanat. *Programme d'Amélioration de La Qualité, de*

References

l'Equité et de La Transparence (PAQUET): Secteur Education-Formation 2013–2025. https://planipolis.iiep.unesco.org/sites/default/files/ressources/senegal-education-sector-plan-2013-2025.pdf.

— — — —. 2017. Ministère de l'Éducation nationale et de l'Alphabétisation. Service National de Coordination de l'Evaluation du Système Éducatif. *Evaluation Des Acquis Des Élèves Du CP2 En Français et Calcul: Rapport Technique*. Dakar.

— — — —. 2018. *Lettre de Politique Générale Pour Le Secteur de l'éducation et de La Formation (LPGS-EF).*

Senegal (PADES programme); AFD (Agence Française de Développement). 2018. *Document de Présentation du Programme d'Appui au Développement de l'Éducation Au Sénégal - PADES 2019–2023*. https://www.education.sn/sites/default/files/2020-01/document%20de%20programme%20PADES%20Sénégal-AFD-PME_QAR2-dec18%20%281%29.pdf

Senghor, O. 2014. *Disseminating and Using Student Assessment Information in The Gambia*. Systems Approach for Better Education Results (SABER) student assessment working paper, 11. Washington, DC. World Bank. https://openknowledge.worldbank.org/handle/10986/17542.

Sutcliffe, S.; Court, J. 2005. *Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries?* Overseas Development Institute. https://cdn.odi.org/media/documents/3683.pdf.

TALENT (Teaching and Learning Educators' Network for Transformation). 2020. *Effective Use of Large-Scale Learning Assessments to Improve Learning Outcomes and Education System Reform*. SDG4 Education 2030 in sub-Saharan Africa Policy Brief. Dakar: UNESCO Office Dakar. https://unesdoc.unesco.org/ark:/48223/pf0000375333/.

References

— — — —. n.d. *Effective Use of Large-Scale Learning Assessments to Improve National Education Outcomes* [Draft Paper]. Policy brief. TALENT: Teaching and Learning Educators' Network for Transformation.

Tobin, M.; Lietz, P.; Nugroho, D.; Vivekanandan, R.; Nyamkhuu, T. 2015. *Using Large-Scale Assessments of Students' Learning to Inform Education Policy. Insights from the Asia-Pacific Region*. Melbourne and Bangkok: ACER and UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000235469.

Tobin, M.; Nugroho, D.; Lietz, P. 2016. 'Large-Scale Assessments of Students' Learning and Education Policy: Synthetising evidence across world regions'. In: *Research Papers in Education,* 31 (5), 578–594. https://doi.org/10.1080/02671522.2016.1225353.

TTF (International Task Force on Teachers for Education 2030) and UNESCO. 2019. *Teacher Policy Development Guide*. Paris: TTF and UNESCO.

UIS (UNESCO Institute for Statistics). 2012. *International Standard Classification of Education, ISCED 2011*. Montreal: UIS. https://unesdoc.unesco.org/ark:/48223/pf0000219109.

— — — —. 2015. *Guinée - Évaluation des Acquis des Élèves en Calcul et en Français*. Central Data Catalog. Montreal: UIS. http://nada.uis.unesco.org/nada/en/index.php/catalogue/157.

— — — —. 2018a. *The Impact of Large-Scale Learning Assessments*. Montreal: UIS. http://uis.unesco.org/sites/default/files/documents/impact-large-scale-assessments-2018-en.pdf.

— — — —. 2018b. *SDG 4 Data Digest: Data to nurture learning*. Montreal: UIS. https://unesdoc.unesco.org/ark:/48223/pf0000366287.

References

— — — —. 2019. *Catalogue of Learning Assessments 2.0*. http://gaml.uis.unesco.org/cla/.

'UIS Glossary'. n.d. http://uis.unesco.org/en/glossary.

— — — —. 2017. *More than Half of Children and Adolescents Are Not Learning Worldwide*. Factsheet 46. Montreal: UIS. http://uis.unesco.org/sites/default/files/documents/fs46-more-than-half-children-not-learning-en-2017.pdf.

UNESCO. 2019. *The Promise of Large-Scale Learning Assessments: Acknowledging limits to unlock opportunities.* Paris: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000369697.

'UNESCO Thesaurus'. n.d. http://vocabularies.unesco.org/browser/thesaurus/fr/page/?clang=en&uri=concept4877.

UNICEF. n.d. *Reaching Educationally Marginalized Children: Namibia. Brief*. https://www.moe.gov.na/files/downloads/1a9_Reaching%20educationally%20marginalized%20children.pdf.

Universalia. 2019. Summative GPE Country Program Evaluation: Batch 5, Country 13: Zambia. Montreal: Universalia. https://www.globalpartnership.org/sites/default/files/2019-08-summative-gpe-country-program-evaluation-for-zambia.pdf.

USAID. 2013. *Senegal's Reading Program: PALME. A G2G Case study*. USAID Office of Education. Washington DC: USAID.

— — — —. 2016. *Ghana 2016 National Education Assessment: Report of Findings*. Available at: https://sapghana.com/data/documents/2016-NEA-Findings-Report_17Nov2016_Public-FINAL.pdf

Verger, A.; Novelli, M.; Altinyelken, H.K. (eds). 2017. *Global Education Policy and International Development: New agendas,*

References

issues and policies. 2nd edn. London ; New York: Bloomsbury Publishing.

Verger, A.; Parcerisa, L.; Fontdevila, C. 2019. 'The Growth and Spread of Large-Scale Assessments and Test-Based Accountabilities: A political sociology of global education reforms'. In: *Educational Review,* 71 (1), 5–30. https://doi.org/10.1080/00131911.2019.1522045.

Wagner, D.A. 2011. *Smaller, Quicker, Cheaper: Improving learning assessments for developing countries*. Paris: IIEP-UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000213663.

Wagner, D.A.; Lockheed, M.; Mullis, I.; Martin, M.O.; Kanjee, A.; Gove, A.; Dowd, A.J. 2012. 'The Debate on Learning Assessments in Developing Countries'. In: *Compare: A Journal of Comparative and International Education,* 42 (3), 509–545. https://doi.org/10.1080/03057925.2012.670480.

Wagner, D.A.; Wolf, S.; Boruch, R. 2018. *Learning at the Bottom of the Pyramid: Science, measurement, and policy in low-income countries*. Paris: IIEP-UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000265581.

Wako, T.N. 2003. *Education Management Information Systems (EMIS): A guide for young managers*. Harare: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000220621.

World Bank. 2009. *Zambia Student Assessment: SABER Country Report 2009*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/28133.

————. 2011. *The Gambia Education Country Status Report*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/2799#:~:text=Although%20much%20has%20been%20achieved,and%20learning%20outcomes%20are%20weak.

References

———. 2013. *Ghana Student Assessment: SABER Country Report 2013*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/20158.

———. 2017. *The Gambia: Education Sector Public Expenditure Review: An efficiency, effectiveness, equity, adequacy, and sustainability analysis*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/28942.

———. 2018. *World Development Report 2018: Learning to realize education's promise*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/28340.

———. 2019a. *Ending Learning Poverty: What will it take?* Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/32553.

———. 2019b. *Ghana Accountability for Learning Outcomes Project: Project Information Document (PID)*. Washington, DC: World Bank.

Zambia. 2005. *Educating the Nation: Strategic Framework for Implementation of Education for All*. https://planipolis.iiep.unesco.org/sites/default/files/ressources/zambia_educ_nation.pdf.

———. 2015. Ministry of General Education and Examinations Council of Zambia. *Zambia's National Assessment Survey Report 2014: Learning Achievement at the Primary School Level*. Lusaka.

———. 2017. Ministry of General Education. *National Learning Assessment Framework (NLAF)*. https://planipolis.iiep.unesco.org/sites/default/files/ressources/zambia_national_learning_assessment_frameworknlaf.pdf.

———. 2019a. Examinations Council of Zambia. *Zambia Education Enhancement Project (ZEEP): Literature Review of*

References

Best Practices to Improve Learner Performance in Zambian Schools. Lusaka: Examinations Council of Zambia.

———. 2019b. Ministry of General Education. *Revised Harmonized Data Collection Instruments (Version 1.0).*

In the same series

2021

Policy Briefs: The Use of Learning Assessment Data

The Gambia
Ghana
Guinea
Namibia
Senegal (in French)
Zambia

Information Sheets: The Use of Learning Assessment Data

The Gambia
Ghana
Guinea (in French)
Namibia
Senegal (in French)
Zambia

Country case studies

Utilisation des données d'évaluation des apprentissages en Guinée (in French)

2019

Working Paper

Use of Learning Assessment Data in Education Policy-Making

About the book

Although the availability of learning data in sub-Saharan Africa has grown, this has not always been followed by their effective use in educational planning. To better understand the reasons for this, IIEP-UNESCO launched a research programme to explore the following questions.

How do countries in the region use data from large-scale learning assessments in different planning cycle phases? What facilitates and impedes this use? What regulatory frameworks are in place? How can governments and development partners sustainably improve the use of learning data?

Collecting IIEP's research findings from The Gambia, Ghana, Guinea, Namibia, Senegal, and Zambia, this publication explores the complex dynamics of the use of learning data, examining how actors' interactions, as well as other factors, play out. It provides a comprehensive analysis on how to improve the use of learning data in sub-Saharan Africa.

About the authors

Ieva Raudonytė is an Education Researcher with the UNICEF Office of Research (Innocenti) working on the Data Must Speak (DMS) Positive Deviance research. Prior to this, she was an Associate Research Officer in the IIEP Research and Development team, where she contributed to research activities related to the theme of teaching and learning. She coordinated the IIEP research programme on the use of learning assessment data and contributed to IIEP's Learning Portal activities.

Tuamanaia Foimapafisi is an education consultant. She contributed to the implementation of the IIEP research programme on the use of learning assessment data. Prior to that, she was a research assistant in a study on racial discrimination in New Caledonia.

unesco
International Institute for Educational Planning

ISBN: 978-92-803-1453-3