



AUGUST 2026

Assessment

From Assessment Data to Better Evidence

Strengthening Africa's Foundational
Learning Evidence Ecosystem

Authors:

Cally Ardington, Sphiwe Bogatsu, Mikaela Daries
and Chifundo Kanjala (AFLEARN, University of Cape Town)



African Foundational Learning Data Hub

DataFirst

University of Cape Town

www.datafirst.uct.ac.za/aflearn



About us

Building Capacity for Educational Change

The hub's goal is to improve learning outcomes for children across the continent by supporting the effective use of foundational learning data for evidence-based advocacy and policy making.

We aim to:

- Create an **African hub of foundational learning data** and expertise in methods, measurement, data quality and comparability.
- **Empower African organisations and ministries to utilise such data** for evidence-based advocacy and policymaking.
- Support an **African research agenda by expanding and capacitating the group of researchers** working on foundational learning.



Contents

Executive Summary	6
Mapping Africa's foundational learning assessment landscape	6
Strengthening the data value chain	6
Demonstrating the value of open microdata	7
A central message	9
Chapter 1: From Assessment to Evidence	10
1.1 A decade of progress and continuing challenges in measuring foundational learning	10
1.2 Building a foundational learning evidence ecosystem	10
1.3 From assessment to evidence	12
1.4 Purpose and scope of this report	13
Chapter 2: Africa's Foundational Learning Assessment Landscape	14
2.1 Introduction	14
2.2 The major sources of nationally representative foundational learning data	14
2.3 Persistent gaps and challenges	17
2.4 Summary and looking ahead	18
Chapter 3: From Data to Evidence: Building Africa's Foundational Learning Data Ecosystem	19
3.1 Assessment data are long-term public assets	19
3.2 Why microdata matter	19
3.3 The current landscape of foundational learning data access	21
3.4 Building a FAIR data ecosystem	22
3.5 Strengthening Africa's foundational learning data ecosystem	22
Chapter 4: From Data to Evidence: Using Household Assessment Microdata to Strengthen the Interpretation and Comparability of Foundational Learning Assessments	24
4.1 Introduction	24
4.2 Household assessment microdata as a resource for comparative research	25
Question 1. Who is represented in the data?	27
Question 2. What do headline indicators hide?	34
Question 3. Which inequalities become visible?	39
Question 4. When are comparisons between countries valid?	45
Question 5. When are language comparisons meaningful?	53
4.3 Chapter conclusion	58



Chapter 5: From Assessment Data to Better Evidence	59
5.1 From assessment data to better evidence	59
5.2 Strengthening the evidence ecosystem	60
5.3 The role of AFLEARN	62
5.4 Conclusion	62
References	63
Data sources	64



Executive Summary

Over the past two decades, Africa has experienced a remarkable expansion in the measurement of foundational learning. National, regional and international assessment programmes now provide an increasingly rich picture of children's reading and numeracy across the continent. These investments have transformed our ability to monitor learning, identify educational challenges and evaluate progress towards national and global goals.

At the same time, the rapid growth of assessment has created a new challenge. The value of these investments depends not only on collecting high-quality data, but also on ensuring that the resulting data remain available, well documented and actively used. Too often, assessment datasets remain difficult to locate, insufficiently documented, inaccessible to researchers or at risk of being permanently lost. As a result, much of the potential value of Africa's investment in foundational learning assessment remains unrealised.

This report argues that strengthening Africa's foundational learning evidence ecosystem offers one of the highest-return investments that governments, assessment organisations and development partners can make. It presents a vision in which assessment data are treated not as one-off project outputs but as long-term public assets that continue to generate evidence, improve future assessments and inform educational policy long after fieldwork has been completed.

The report is organised around four interconnected themes.

Mapping Africa's foundational learning assessment landscape

The report first documents the rapid expansion of foundational learning assessment across Africa. National, regional and international programmes now cover much of the continent and employ a wide range of assessment approaches, languages and sampling designs. Together they represent one of the richest foundational learning assessment landscapes in the world.

This expansion has greatly improved the availability of evidence on children's learning, but important challenges remain. Coverage remains uneven across countries and languages, assessment designs differ substantially, and comparability cannot be assumed. The report argues that these differences should be viewed not as weaknesses but as opportunities for complementary evidence, provided that researchers understand the characteristics of each programme and the populations they represent.

Strengthening the data value chain

The report then examines what happens to assessment data after they have been collected. In recent years, access to foundational learning microdata has considerably improved. This is evident from the recent public release of PASEC microdata, the publication of nationally representative ICAN-ICAR dataset, and the recovery of more than 20 years of EGRA datasets. Despite this, access remains highly fragmented across assessment programmes.

Many nationally commissioned assessments remain unavailable for secondary analysis, documentation standards vary widely and valuable historical datasets continue to face the risk of permanent loss.

The report proposes the FAIR principles (Findable, Accessible, Interoperable and Reusable) as a practical framework that strengthens Africa's foundational learning data ecosystem. Applied to assessment data, these principles increase transparency, support reproducible research, and unlock analyses the original collectors never foresaw



Demonstrating the value of open microdata

The central contribution of the report is to demonstrate what becomes possible when publicly available microdata are combined with rigorous secondary analysis. The report explores five methodological questions using nationally representative household assessment data from UNICEF's Multiple Indicator Cluster Surveys (MICS) and the PAL Network's International Common Assessment of Numeracy and Reading (ICAN-ICAR).

First, it shows that understanding who is represented in assessment data requires more than examining sampling designs. Household assessment microdata allow researchers to investigate school participation, school sector coverage, realised response rates and assessment eligibility, providing a much clearer understanding of the populations represented by different assessments.

Second, it demonstrates that headline indicators such as average scores and minimum proficiency rates necessarily simplify children's learning. Examining score distributions and individual assessment components reveals important differences that remain hidden when analyses rely solely on national averages or proficiency benchmarks.

Third, the report illustrates how linking learning outcomes to household characteristics reveals patterns of educational inequality that cannot be observed in published reports alone. Household assessment microdata make it possible to investigate wealth gradients, school participation and children's educational pathways in far greater detail.

Fourth, the report argues that international comparisons should be investigated rather than assumed. Publicly available microdata enable researchers to examine assessment timing, instructional exposure, age eligibility and learning trajectories, thereby strengthening the interpretation of cross-country comparisons.

Finally, the report explores how children's language environments influence assessment results. By linking information on home language, language of instruction and assessment language, household surveys provide new opportunities to investigate one of the most important challenges facing comparative foundational learning assessment in Africa.

Together, these examples demonstrate that open microdata improve not only transparency but also the interpretation, comparability and quality of assessment evidence itself.

Building Africa's foundational learning evidence ecosystem

The report concludes by arguing that Africa's next challenge is not simply to collect more assessment data, but to maximise the value of the substantial investments already made.

This requires strengthening every stage of the evidence ecosystem: from assessment design and documentation through preservation, open access, capacity building and methodological research. Governments, assessment organisations, universities, regional bodies and development partners each have distinct but complementary roles to play.

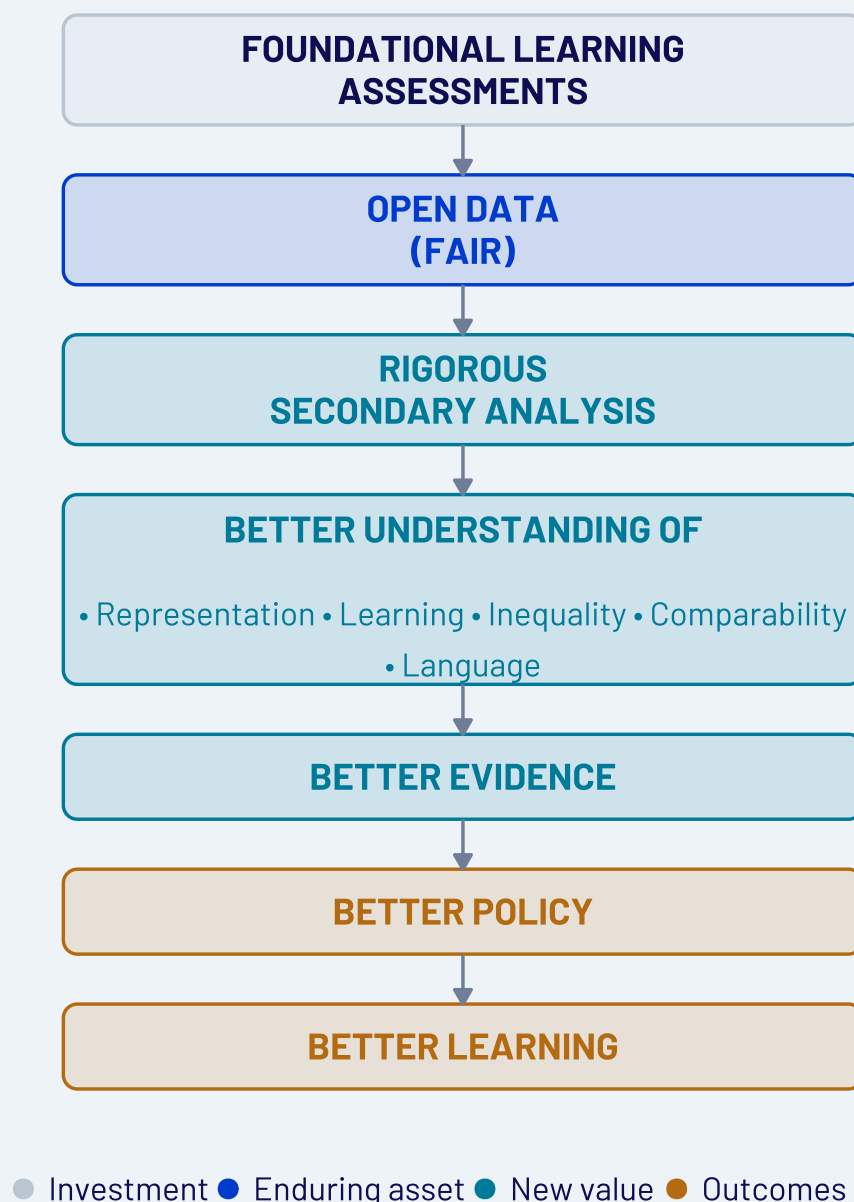
The report identifies six priorities for the coming decade:

- Design assessments with long-term data reuse in mind.
- Preserve and document existing assessment datasets before valuable evidence is lost.
- Make assessment microdata as open as possible through FAIR-aligned governance and trusted repositories.
- Build African capacity for secondary analysis through training, analytical resources and open methods.
- Support methodological research that improves assessment quality, comparability and interpretation.
- Strengthen collaboration across institutions responsible for generating, stewarding and using foundational learning evidence.



Open microdata make assessment evidence more useful and more reliable

Strengthening transparency, interpretation, comparability and assessment quality



The African Foundational Learning Data Hub (AFLEARN) supports these priorities by strengthening data stewardship, harmonising documentation, improving access to assessment data, developing analytical resources, building research capacity and advancing methodological work on foundational learning assessment across the continent.



A central message

Africa has already built one of the world's richest and most diverse foundational learning assessment landscapes. The next frontier is to ensure that these assessments continue to generate value long after data collection has ended.

Open, well-documented and reusable assessment microdata transform data into evidence. They increase the return on investments in assessment, strengthen transparency and comparability, support methodological innovation and enable richer analyses of children's learning than published reports alone.

Collecting assessment data is only the beginning of the evidence journey. By treating assessment data as long-term public assets, Africa can transform its substantial investment in foundational learning assessment into a lasting foundation for better research, better policy and, ultimately, better learning opportunities for every child.





Chapter 1: From Assessment to Evidence

1.1 A decade of progress and continuing challenges in measuring foundational learning

Foundational learning—the ability to read with understanding and apply basic numeracy—is the foundation upon which all subsequent learning is built. Children who do not acquire these skills during the early years of schooling face increasing difficulty accessing the curriculum, limiting both their educational opportunities and their future life chances. For this reason, foundational learning has become one of the defining education priorities of the Sustainable Development Goal (SDG) era and an increasingly important focus of education policy across Africa.

Over the past decade, the measurement of foundational learning has expanded rapidly across the continent. Governments, regional organisations, development partners, universities and civil society organisations have invested heavily in developing assessments capable of measuring whether children are acquiring the reading and numeracy skills expected during the early years of schooling. National assessment systems have been strengthened, regional programmes have expanded their coverage, and new household-based assessments have made it possible to measure learning among children who are often invisible in school-based data.

This expansion has fundamentally transformed the evidence available to policymakers. Fifteen years ago, nationally representative information on foundational learning was available for only a relatively small number of African countries. Today, most countries have participated in at least one national, regional or international assessment, and many have accumulated multiple sources of evidence collected over time. These assessments have revealed that millions of African children complete the early years of schooling without acquiring foundational reading and numeracy skills. They have also highlighted large inequalities within and between countries and shifted policy discussions from school participation alone to the quality of learning taking place in classrooms.

This progress has been driven by several complementary developments. The adoption of Sustainable Development Goal 4 created an urgent demand for internationally comparable measures of learning, particularly Indicator 4.1.1(a), which monitors the proportion of children achieving minimum proficiency in reading and mathematics by the end of lower primary school. At the same time, governments increasingly recognised that expanding access to education was not sufficient if children were not learning. Attention therefore shifted from measuring schooling to measuring learning.

Despite these advances, Africa's foundational learning assessment landscape remains a work in progress. Coverage continues to vary across countries, assessment programmes differ in their objectives, designs and target populations. Thus, ensuring meaningful comparability across assessments and over time remains a major methodological challenge. Many national assessment systems continue to evolve, while questions surrounding language of assessment, measurement frameworks, linking across assessment cycles and the interpretation of learning indicators remain active areas of research. Continued investment in high-quality assessment therefore remains essential.

At the same time, the continent has accumulated a substantial body of assessment data whose value extends well beyond the publication of national reports. Increasingly, the challenge is not only to generate new evidence but also to maximise the value of the evidence that already exists.

1.2 Building a foundational learning evidence ecosystem

The rapid expansion of foundational learning assessment has created more than a growing collection of datasets. It has given rise to an increasingly sophisticated evidence ecosystem in which governments, assessment programmes, researchers, universities, repositories, data infrastructure organisations, civil



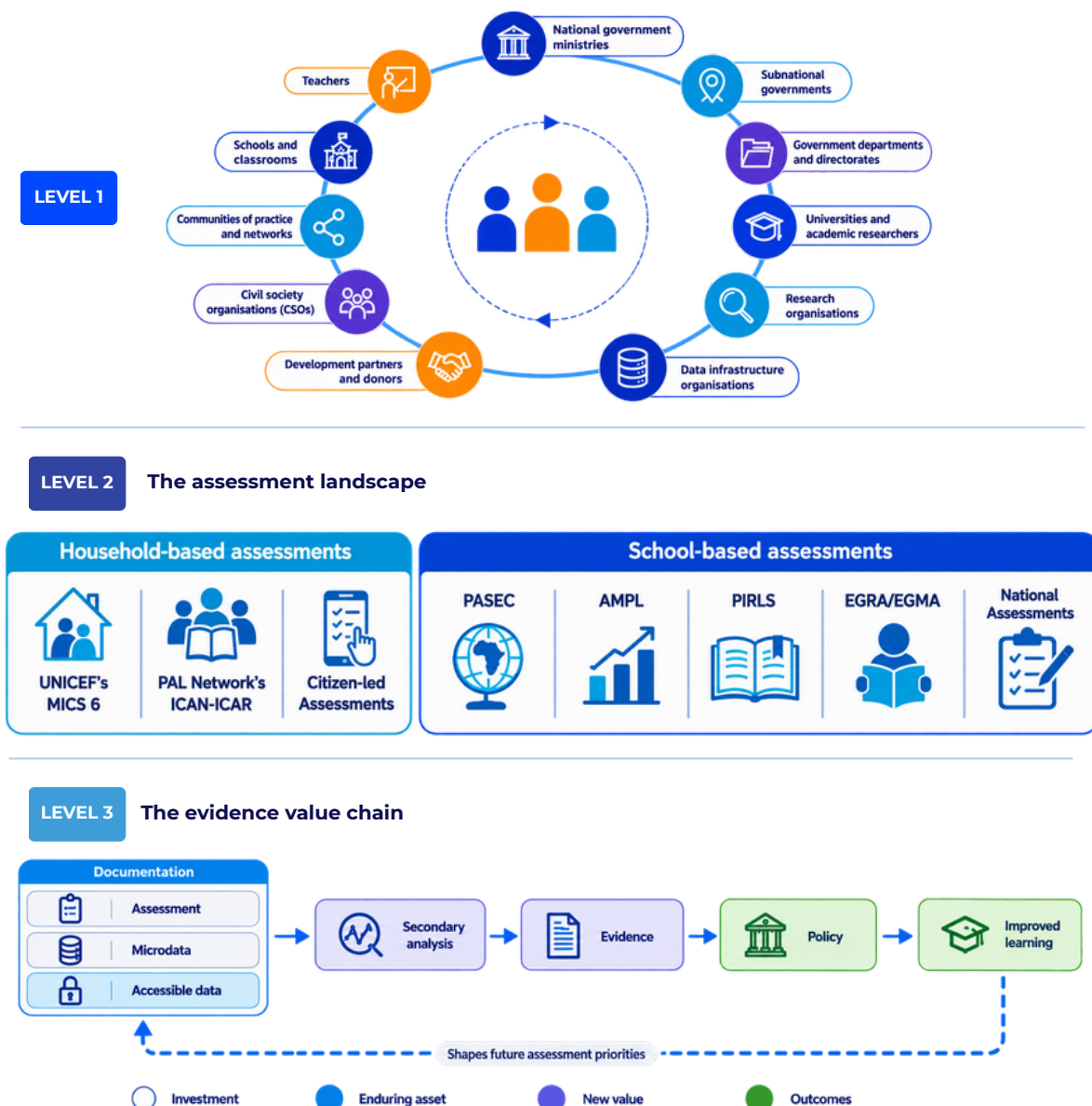
society organisations and development partners collectively generate, curate, analyse and use evidence to improve learning.

Thinking in terms of an evidence ecosystem shifts attention away from individual assessments towards the broader processes through which evidence creates public value. High-quality assessments are an essential starting point, but they do not by themselves improve education systems. Their value depends equally on whether assessment data are preserved, documented, made accessible, analysed rigorously and translated into evidence that informs policy and practice.

Figure 1.1 presents the conceptual framework that underpins this report by illustrating three complementary perspectives on Africa's foundational learning evidence ecosystem.

FIGURE 1

The foundation learning evidence ecosystem



Level 1 represents the Foundational Learning Evidence Ecosystem. Organised as a network rather than a linear process, it illustrates the diverse actors involved in generating and using evidence. Governments establish assessment policies and invest in data collection. Assessment programmes generate learning data.



Universities and researchers develop new analytical methods and produce evidence. Data infrastructure organisations curate, preserve and disseminate datasets. Civil society organisations, development partners and policymakers translate evidence into action. Together these actors form an interconnected system in which evidence is continuously generated, interpreted and refined.

Level 2 represents the Assessment Landscape—the data-generating component of the ecosystem. National, regional and international assessment programmes each contribute complementary evidence while differing in their objectives, populations, methodologies and measurement frameworks. No single assessment can answer every policy question. Rather, it is the combination of these programmes that provides an increasingly rich picture of foundational learning across Africa. Chapter 2 examines this landscape in detail.

Level 3 represents the Evidence Value Chain, illustrating how assessment data increase in value over time. The process begins with the conceptualisation, design and implementation of an assessment, but it does not end with data collection or the publication of a national report. Rather, value continues to be created through data curation, documentation, dissemination, discovery, secondary analysis and evidence synthesis. Each stage broadens the range of questions that can be answered and increases the return on investments in assessment. Chapters 3 and 4 examine this value chain in greater depth.

These three levels should not be viewed as sequential, but as complementary perspectives on the same system. Together they demonstrate that strengthening Africa's foundational learning evidence base requires investment not only in assessment programmes themselves, but also in the institutions and processes that enable assessment data to generate value long after fieldwork has been completed.

1.3 From assessment to evidence

Continued investment in high-quality assessment remains fundamental to improving foundational learning across Africa. Many countries continue to strengthen national assessment systems, expand coverage, improve measurement frameworks and address important challenges relating to comparability, language and inclusion. These investments remain essential.

Alongside these continuing developments, however, a complementary opportunity has emerged. The continent has accumulated a substantial body of nationally representative assessment data that can support far more than the analyses presented in published assessment reports.

National reports necessarily address only a limited number of policy questions. They typically present national averages, selected subgroup comparisons and headline indicators designed to support monitoring and accountability. While these analyses are indispensable, they however only represent a small fraction of the knowledge contained within the underlying assessment microdata.

When assessment datasets are accompanied by high-quality metadata and clear access arrangements, they become long-term public assets rather than one-off project outputs. Publicly available microdata allow researchers to investigate educational inequalities, evaluate the quality and comparability of assessments, develop new analytical methods, replicate published findings and answer questions that were never envisaged during the original design of the assessment.

Realising this potential requires more than simply releasing datasets. Assessment data must be accompanied by sufficient structured metadata to enable discoverability, lossless exchange and meaningful reuse. Researchers require analytical capacity, supporting resources and opportunities to share methods and findings. Data infrastructure organisations play a crucial role in preserving datasets and ensuring that they remain accessible long after individual projects have ended. Together, these activities transform assessment data into an enduring public resource.

The evolution of Africa's foundational learning evidence ecosystem therefore requires complementary investments in three areas: continuing to strengthen assessment systems; improving the stewardship and curation of assessment data; and expanding the rigorous secondary analysis that transforms data into new



knowledge. These investments reinforce one another. Better assessments generate better data, while wider access and stronger analysis improve both the interpretation of existing assessments and the design of future ones.

1.4 Purpose and scope of this report

This report is about strengthening Africa's foundational learning evidence ecosystem through improved access to, and use of, assessment microdata. It argues that maximising the value of investments in foundational learning requires attention not only to how assessments are designed and implemented, but also to how the resulting microdata are stewarded, shared and used.

The report has four objectives.

First, it reviews the current landscape of foundational learning assessments across Africa, highlighting the complementary roles played by national, regional and international assessment programmes and identifying the strengths and limitations of the existing evidence base.

Second, it examines the current state of data access and stewardship. The report reviews which assessment datasets are publicly available, where they can be accessed, and the barriers that continue to limit their discoverability, accessibility and reuse. It argues that adopting the principles of Findability, Accessibility, Interoperability and Reusability (FAIR) would substantially increase the value generated by existing investments in learning assessment.

Third, and most importantly, the report demonstrates how publicly available microdata can be used to strengthen the interpretation, quality and comparability of foundational learning assessments. Using two nationally representative household-based assessment programmes—UNICEF's Multiple Indicator Cluster Surveys (MICS) Foundational Learning Skills module and the PAL Network's International Common Assessment of Numeracy and Reading (ICAN-ICAR)—the report illustrates how secondary analysis can address methodological questions that cannot be answered using published reports alone. These include questions relating to representation, proficiency benchmarks, learning distributions, instructional exposure, socioeconomic inequality, age eligibility and language.

Finally, the report argues that publicly available assessment data are more than a resource for researchers. They are an essential component of a broader evidence ecosystem that supports methodological innovation, transparency, capacity strengthening and evidence-informed policymaking across Africa.

The report situates AFLEARN and its contribution to improving data access and analytical capacity within a broader emerging conversation about the future of educational evidence on the continent. It argues that assessment data should increasingly be viewed as long-term public assets whose value depends not only on how well they are collected, but also on how effectively they are curated, shared, interpreted and reused.

Key takeaway. Africa has made remarkable progress in measuring foundational learning, but important challenges remain in expanding coverage, strengthening comparability and improving assessment quality. At the same time, the continent has accumulated a substantial body of assessment data whose value extends far beyond the publication of national reports. Maximising the return on these investments requires continued improvements in both assessment systems and the broader evidence ecosystem—including data stewardship, public access and rigorous secondary analysis. The following chapters examine how these complementary investments can strengthen the interpretation, comparability and policy value of foundational learning assessments across Africa.



Chapter 2: Africa's Foundational Learning Assessment Landscape

2.1 Introduction

This chapter provides an overview of the principal assessment programmes that generate foundational learning data across the continent.

Africa's assessment landscape has expanded substantially over the past two decades. Governments, regional organisations, development partners and research institutions have invested in a wide range of national, regional and international assessments to better understand whether children are acquiring the reading and numeracy skills expected during the early years of schooling. These programmes differ in their objectives, populations, assessment frameworks and intended uses, but together they have created an increasingly rich—although still incomplete—body of evidence on foundational learning across the continent.

No single assessment can answer every policy question. Some programmes are designed to monitor national education systems, others to facilitate international comparisons, while household-based assessments provide insights into the learning of children who may never appear in school-based datasets. Understanding these complementary roles is more informative than attempting to identify a single "best" assessment.

The purpose of this chapter is therefore not to evaluate individual assessment programmes, but to introduce the principal sources of nationally representative foundational learning data available in Africa. These assessments form the data-generating component of the evidence ecosystem introduced in Chapter 1 and provide the foundation for the discussions of data access in Chapter 3 and secondary analysis in Chapter 4.

2.2 The major sources of nationally representative foundational learning data

At the broadest level, Africa's foundational learning assessments fall into two groups: household-based assessments, which sample children directly from households, and school-based assessments, which sample children through schools. Within each of these groups are national, regional and international programmes designed to answer different policy questions.

Figures 2.1 and 2.2 summarise the current geographic coverage of the principal nationally representative household- and school-based assessment programmes across Africa. Together they illustrate both the remarkable expansion of foundational learning assessment and the diversity of approaches that have emerged.

Household-based assessments

Household-based assessments measure learning among children in the population rather than only those attending school. This allows learning outcomes to be analysed alongside household characteristics and, importantly, includes children who are out of school as well as those who are enrolled.

An important strand of household-based learning measurement in Africa has been the Citizen-Led Assessment (CLA) movement. Adapted from India's ASER model, CLAs have been implemented by civil-society organisations in a number of African countries, including through Uwezo in Kenya, Tanzania and Uganda, Jàngandoo in Senegal, Bèekunko in Mali and LEARNigeria. These assessments typically use short, orally administered tasks in reading and numeracy and are conducted in households rather than schools. Their household-based design enables them to capture learning among children regardless of school enrolment or attendance, while their relatively simple tools and large-scale implementation have also been used to stimulate public engagement around foundational learning. The CLA movement demonstrated that learning



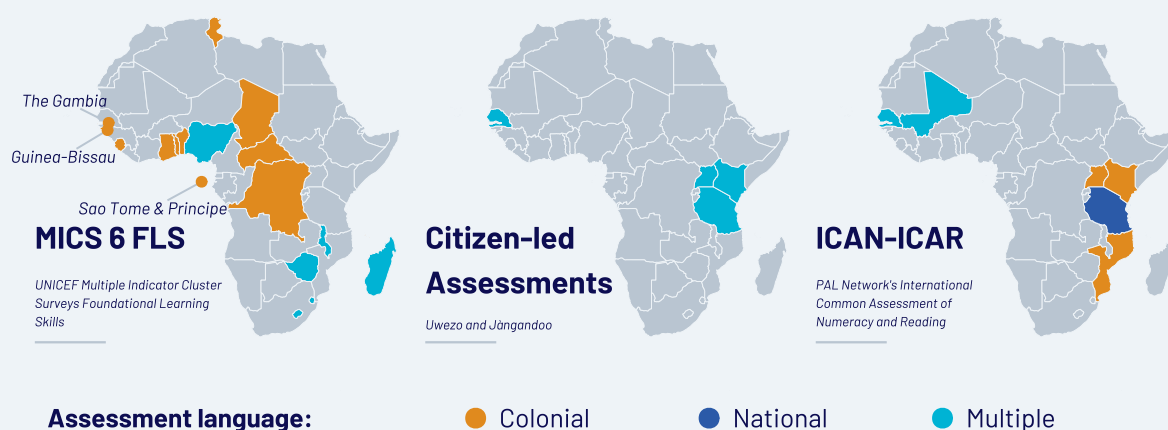
could be measured at population level outside the formal school system and laid important foundations for subsequent household-based assessment initiatives in Africa.

Building on this approach, the International Common Assessment of Numeracy and Reading (ICAN-ICAR) was developed by the People's Action for Learning (PAL) Network to provide a more standardised framework for measuring foundational learning across countries. ICAN and ICAR assess children aged 5–16 years in households and combine nationally representative sampling with common assessment frameworks and psychometric methods designed to support both national reporting and cross-country comparison. The first nationally representative African ICAN-ICAR implementation was conducted in 2025 across twelve countries, creating an important new source of publicly available household assessment data. ICAN-ICAR is discussed in detail in Chapter 4.

FIGURE 2.1

Household-based assessments have expanded learning measurement across Africa

Nationally representative household-based foundational learning assessments in Africa.



A second major source of household-based learning data is the Multiple Indicator Cluster Surveys (MICS) Foundational Learning Skills module, coordinated by UNICEF. MICS embeds direct assessments of reading and numeracy within a nationally representative household survey. Children aged 7–14 years complete short assessments of foundational reading and mathematics alongside the broader household questionnaire, allowing learning outcomes to be analysed together with detailed information on household characteristics, education, health and child wellbeing. Because MICS is implemented across a large number of countries using a common survey platform, it has become one of the largest sources of comparable household-based learning data in Africa. MICS is also discussed in detail in Chapter 4.

School-based assessments

School-based assessments remain the principal source of nationally representative evidence on learning within education systems. They provide detailed information about learning among children attending school and are typically accompanied by extensive contextual information on schools, teachers and classrooms.

The major school-based assessment programmes include:

Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN (PASEC). Coordinated by La Conférence des Ministres de l'Éducation des pays ayant le français en partage (CONFEMEN), PASEC assesses learning in Grades 2 and 6 across Francophone Africa using common assessment frameworks and nationally representative samples. Since 2014 the programme has operated through harmonised regional assessment cycles, providing one of the continent's richest sources of internationally comparable school-based learning data.

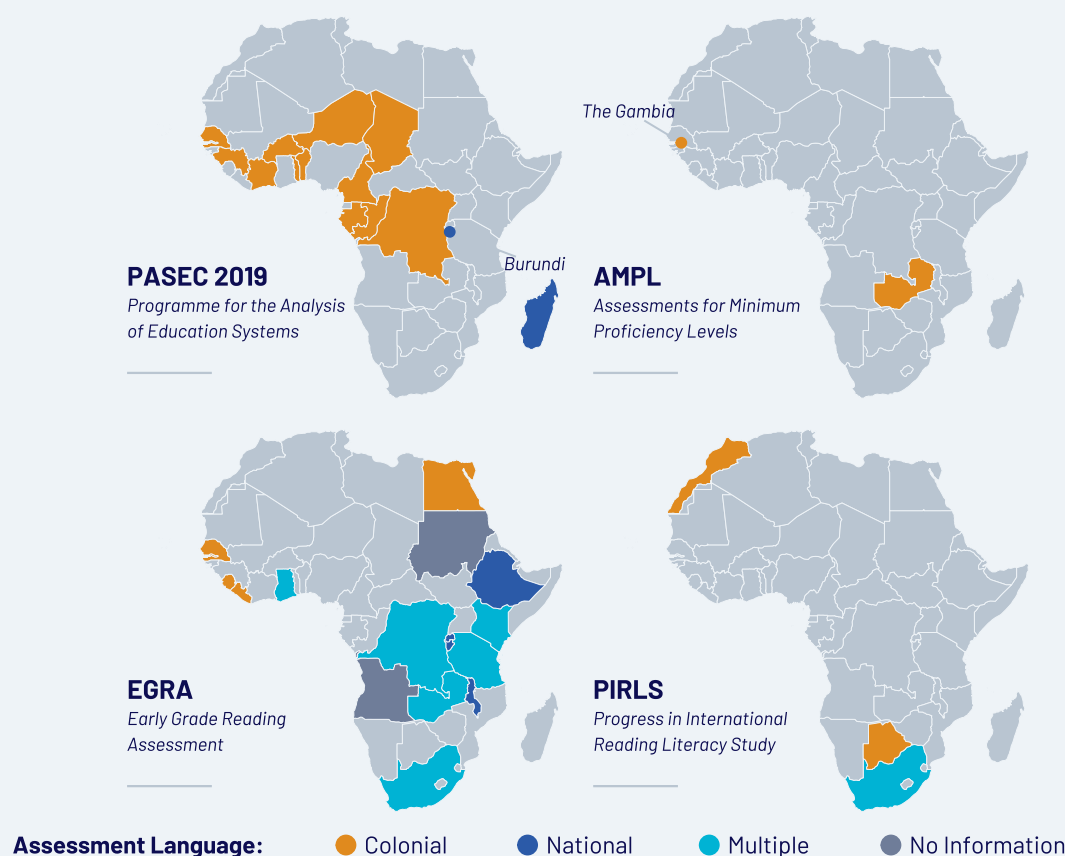
Assessments for Minimum Proficiency Levels (AMPL). Developed by the UNESCO Institute for Statistics and the Australian Council for Educational Research, AMPL was designed specifically to support reporting against Sustainable Development Goal Indicator 4.1.1(a). The assessment has been implemented in a small number of African countries and provides an important reference point for internationally comparable measurement of foundational learning.

Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA). EGRA and EGMA are among the most widely implemented foundational learning assessments in Africa. Rather than representing a single assessment programme, they comprise a flexible assessment toolkit that has been adapted to numerous languages and contexts. This flexibility has facilitated widespread adoption while limiting direct comparability across countries. Implementations vary considerably in sampling design, geographical coverage and assessment content, and many are not nationally representative.

FIGURE 2.2

School-based assessments provide broad but uneven coverage of foundational learning across Africa

Nationally representative school-based foundational learning assessments in Africa



Progress in International Reading Literacy Study (PIRLS). PIRLS is a global assessment of Grade 4 reading comprehension conducted every five years. Although relatively few African countries have participated, PIRLS provides an important international benchmark and has contributed substantially to understanding reading achievement in countries such as South Africa, Morocco, Egypt and Botswana.



National assessments. Many African countries also operate their own national assessment programmes. These are typically aligned with national curricula and policy priorities and play a central role in monitoring education systems. Designs vary substantially across countries in terms of grades assessed, frequency of implementation, technical quality and public accessibility, but collectively these assessments represent an essential component of Africa's foundational learning evidence base.

Understanding the assessment landscape

No single assessment programme should be regarded as the definitive source of evidence on foundational learning. Household-based assessments provide information about all children, including those who are out of school. School-based assessments provide richer information about education systems, schools and classrooms. Regional and international assessments facilitate cross-country comparison, while national assessments are often more closely aligned with local curricula and policy priorities. These programmes are therefore best viewed as complementary components of a broader evidence ecosystem rather than competing alternatives.

2.3 Persistent gaps and challenges

The expansion of foundational learning assessment represents remarkable progress, but significant challenges remain. Coverage has increased substantially, yet important gaps persist in representation, comparability, inclusion and sustainability.

Coverage and representation

Although most African countries have now participated in at least one nationally representative assessment, the frequency and continuity of assessment vary considerably. Some countries have accumulated multiple assessment cycles spanning more than a decade, while others have only a single nationally representative dataset. Household and school-based assessments also represent different populations, meaning that no single programme provides a complete picture of children's learning.

Comparability

Differences in assessment design continue to limit direct comparisons across programmes and countries. Assessments differ in their target populations, grades, age ranges, sampling strategies, languages of assessment, measurement frameworks and definitions of minimum proficiency. Temporal comparability also remains challenging, as relatively few programmes have established robust linking across assessment cycles. Improving comparability therefore remains an important priority for the field.

Inclusion

One of the most significant remaining challenges concerns who is excluded from assessment. Children with disabilities continue to be underrepresented in many assessment programmes or are unable to participate meaningfully because assessment instruments have not been adapted to their needs. Encouraging progress has been made through adaptations of EGRA for children with disabilities and through the combination of the MICS Child Functioning Module with the Foundational Learning Skills module, but inclusive assessment remains the exception rather than the norm.

Language presents a further challenge. Many assessments continue to be administered in languages that differ from children's home language or even their language of instruction. As Chapter 4 demonstrates, these choices have important implications for both the validity and comparability of assessment results.



Sustainability

Finally, many assessment programmes remain heavily dependent on external funding. Donor-supported assessments often generate valuable data but are not always repeated once projects conclude. National assessments, while potentially more sustainable, vary considerably in technical capacity, documentation and public accessibility. Strengthening national assessment systems and integrating different sources of learning data into coherent national monitoring frameworks therefore remain important priorities.

2.4 Summary and looking ahead

Collectively, these assessment programmes constitute the data-generating component of Africa's foundational learning evidence ecosystem. Each programme has distinct objectives, populations and methodological strengths, and together they provide complementary perspectives on children's learning across the continent. Understanding this landscape is essential for interpreting the evidence that follows and for recognising that different assessments are designed to answer different questions.

The next chapter shifts attention from what data exist to how those data are stewarded and accessed. It examines the current landscape of public access to foundational learning assessment microdata and argues that improving data curation, documentation and discoverability is essential for maximising the value of Africa's investment in learning assessment. Chapter 4 then demonstrates how publicly available microdata can be used to address methodological questions relating to representation, comparability and interpretation that cannot be answered using published assessment reports alone.



Chapter 3: From Data to Evidence: Building Africa's Foundational Learning Data Ecosystem

3.1 Assessment data are long-term public assets

Chapter 2 demonstrated that Africa has developed an increasingly rich landscape of foundational learning assessments. Together, national, regional and international assessment programmes have generated an extensive body of data describing children's reading and numeracy skills across the continent. These investments represent an enormous achievement and continue to expand as countries strengthen their assessment systems and improve the quality and comparability of learning measurement.

Yet the value of these investments depends not only on the quality of the assessments themselves, but also on what happens after fieldwork has been completed.

Every foundational learning assessment requires substantial investment. Governments allocate scarce public resources to planning assessments, developing instruments, training field teams and collecting high-quality data. Development partners contribute financial resources and technical expertise. Teachers, schools, children and their families devote time to participating in the assessments themselves.

The immediate outputs of these investments are familiar: national reports, policy briefs, presentations and headline indicators. These products are essential. They provide governments with timely evidence on the performance of education systems and support national and international reporting on learning outcomes. However, they represent only one stage in a much longer evidence-generation process.

The most enduring output of an assessment is the underlying microdata.

Unlike published reports, which necessarily answer a predetermined set of questions, assessment microdata continue to generate new knowledge long after the original analyses have been completed. As policy priorities evolve and analytical methods improve, existing datasets frequently become valuable in ways that could never have been anticipated when they were first collected.

Assessment data should therefore increasingly be viewed as **long-term public assets** rather than simply the by-product of an assessment exercise. Like census data, household surveys and administrative records, they are national resources whose value grows through careful documentation, preservation, accessibility and reuse.

Figure 3.1 illustrates this shift in perspective. Assessment should not be viewed as the end of the evidence-generation process but as its beginning. Investment in assessment generates microdata, but these data create relatively little additional value unless they are documented, curated and made accessible. Once this occurs, researchers can undertake secondary analyses that generate new evidence, inform policy and ultimately contribute to improved learning. Importantly, this process is not linear. Insights generated through secondary analysis frequently influence the design of future assessments, creating a virtuous cycle of methodological improvement in which evidence not only informs policy but also strengthens assessment systems themselves.

3.2 Why microdata matter

Published assessment reports typically summarise national learning outcomes and provide disaggregation by characteristics such as gender, region and socioeconomic status. These analyses are indispensable for monitoring learning and informing policy. However, they represent only a small fraction of the information contained within the underlying assessment datasets.

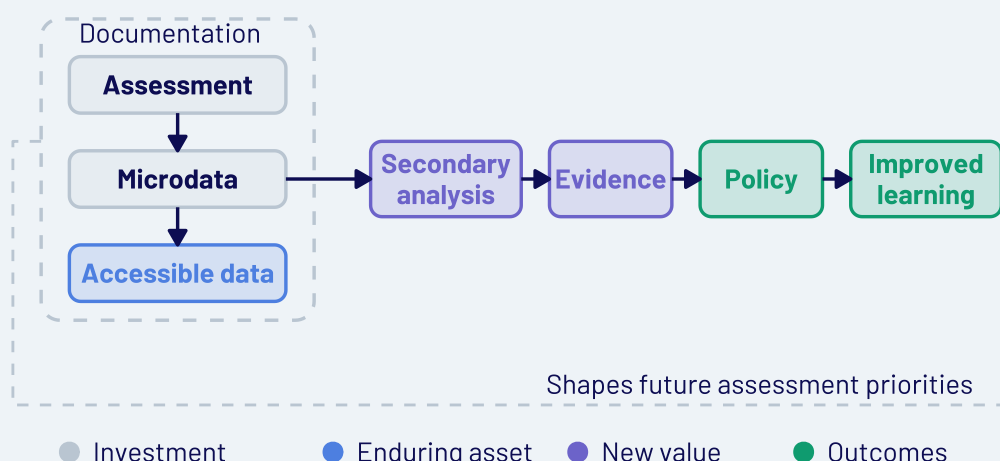
Publicly available microdata make it possible to answer questions that cannot be addressed using published

reports alone. Researchers can examine the complete distribution of learning outcomes rather than relying solely on averages and headline indicators. They can investigate inequalities among different groups of children, analyse performance on individual assessment components, evaluate alternative proficiency measures and study how learning varies across ages, grades and contexts.

FIGURE 3.1

The value of assessment data grows through access, reuse and feedback

The foundational learning data value chain



Further, microdata also allow researchers to investigate methodological questions relating to the quality and comparability of assessments themselves. They make it possible to examine who was represented in the assessment, explore how assessment timing influences comparisons across countries, investigate the consequences of language-of-assessment decisions, evaluate sampling and response patterns, and assess the implications of different measurement frameworks. These are not simply technical issues; they fundamentally influence how assessment results should be interpreted.

Existing datasets also become increasingly valuable as analytical methods evolve. Questions that were not envisaged during the original design of an assessment can often be investigated years later without the financial and operational costs of collecting new data. Public access therefore substantially increases the return on existing investments in assessment.

Accessible microdata consequently:

- broaden the range of questions that can be investigated;
- strengthen transparency through independent replication of published findings;
- improve assessment quality through methodological scrutiny;
- support analytical capacity development;
- encourage methodological innovation; and
- increase the long-term value generated by existing assessments.

Many of these opportunities are illustrated in Chapter 4, which demonstrates how publicly available household assessment microdata can improve the interpretation and comparability of foundational learning assessments.



3.3 The current landscape of foundational learning data access

As documented in Chapter 2, a substantial body of foundational learning assessment data has already been generated across Africa. However, the extent to which these data can be discovered, accessed and reused varies considerably across assessment programmes. While some datasets are accompanied by comprehensive documentation and well-established dissemination mechanisms, others remain difficult to locate, inadequately documented or unavailable for independent analysis. As a result, much of the continent's investment in learning assessment continues to be underutilised.

Table 3.1 summarises the current landscape of public access to the principal foundational learning assessment datasets available in Africa.

TABLE 3.1.

Current state of access to foundational learning assessment microdata

Assessment	Access	Where to access
Household-based assessments		
MICS (FLS)	Public	MICS website: Link IPUMS-MICS (harmonised multi-country datasets): Link
ICAN-ICAR	Public	DataFirst Open Data Portal: Link
Citizen-Led Assessments	Partial	PAL Network datasets page: Link
School-based assessments		
PASEC	Public	CONFEMEN/PASEC Data Portal: Link
AMPL	Public	UNESCO Institute for Statistics: Link
PIRLS	Public	IEA Repository: Link TIMSS & PIRLS International Study Center: Link
EGRA	Partial	AFLEARN Rescued USAID Education Data: Link to map DataFirst AHEAD: Link
National Assessments	Restricted	Usually not publicly accessible

Large international assessment programmes generally provide the strongest examples of open data practice. MICS and PIRLS have long-established systems for disseminating anonymised microdata together with questionnaires, technical reports and supporting documentation. Access to regional assessment data has also improved substantially in recent years. The PASEC microdata are now publicly available through the PASEC Data Portal, while nationally representative ICAN-ICAR datasets are publicly available through AFLEARN at DataFirst. Some earlier Citizen-Led Assessment datasets were previously available through the PAL Network datasets page, but access is uneven and some older links are inactive. Users may need to contact PAL Network member organisations for specific datasets.



The situation is more fragmented for Early Grade Reading Assessments (EGRAs), reflecting the decentralised and project-based nature of their implementation. EGRAs are typically commissioned by donors and implemented by research organisations or consulting firms under short-term contracts. Consequently, responsibility for archiving and disseminating the resulting datasets has often been unclear, with microdata remaining in the possession of implementing agencies or ministries rather than being deposited in public repositories. Although USAID-funded EGRAs were intended to be archived in the USAID Development Data Library, a substantial backlog accumulated over many years. Following the closure of USAID's education office, AFLEARN worked with a range of partners to recover and preserve these data. This initiative resulted in the recovery of more than 2,500 files from 109 assessment projects spanning 43 African countries and over two decades of foundational learning measurement. These rescued datasets are now publicly available through the ICPSR DataLumos platform, substantially expanding public access to one of Africa's most important collections of early grade assessment data.

The greatest gap remains nationally commissioned learning assessments. Even where these assessments are nationally representative and technically robust, their microdata are rarely released beyond the commissioning ministry or assessment agency. This limits opportunities for independent verification, secondary analysis and methodological innovation, while also reducing opportunities for cross-country learning.

3.4 Building a FAIR data ecosystem

Improving access alone is not sufficient. For assessment data to generate long-term value they must also be well documented, easy to discover, interoperable with other datasets and supported by clear governance arrangements.

These principles are captured by the FAIR framework, which encourages data to be:

- **Findable**, through rich metadata and persistent identifiers;
- **Accessible**, through transparent access procedures and sustainable repositories;
- **Interoperable**, using common standards that facilitate integration across datasets; and
- **Reusable**, through comprehensive documentation, clear licensing arrangements and long-term preservation.

Applied to foundational learning assessments, the FAIR principles help move assessment data through the evidence value chain described earlier in this chapter. They ensure that datasets remain useful not only to the organisations that originally collected them, but also to future researchers investigating new questions using new analytical approaches.

Importantly, FAIR does not imply that all datasets should be unrestricted. Rather, data should be as open as possible and as closed as necessary. Appropriate governance remains essential for protecting confidentiality, respecting national ownership and ensuring responsible data use. Public-use, scientific-use and restricted-access datasets each have an important role within a well-functioning data ecosystem.

As access to foundational learning data continues to improve, the use of internationally recognised metadata standards should become standard practice. These standards greatly improve discoverability, facilitate data exchange, and reduce information loss when datasets are disseminated across multiple systems and repositories. Technology already exists. The remaining challenge is to build awareness, institutional capacity and a culture that recognises assessment data as valuable long-term assets rather than short-term project outputs.

Why this matters. Data that cannot be found, understood or reused represent lost opportunities for research, policy and public investment.

3.5 Strengthening Africa's foundational learning data ecosystem

The evidence presented in this chapter suggests that strengthening Africa's foundational learning data ecosystem will require sustained attention to data governance, documentation, preservation, interoperability



and analytical capacity across the institutions that produce, manage and use assessment data. The roles of governments, assessment organisations, universities, development partners, data infrastructure organisations and AFLEARN in advancing this agenda are considered in Chapter 5.

Looking ahead. The principles discussed in this chapter raise an obvious question: beyond making datasets easier to access, what additional value does public access to assessment microdata actually create?

The next chapter answers this question through a series of practical examples using nationally representative household assessment data. Rather than focusing solely on substantive policy findings, it demonstrates how secondary analysis can improve the interpretation, comparability and methodological understanding of foundational learning assessments. In doing so, it illustrates how publicly available microdata transform assessment datasets from static reports into long-term public assets that continue to generate new evidence for research, policy and the improvement of future assessments.

Key takeaway. Collecting assessment data is only the beginning of the evidence-generation process. When assessment microdata are well documented, preserved and made accessible under appropriate governance arrangements, they become long-term public assets that support methodological innovation, strengthen the interpretation and comparability of learning assessments, and generate evidence far beyond the questions addressed in original assessment reports.





Chapter 4: From Data to Evidence: Using Household Assessment Microdata to Strengthen the Interpretation and Comparability of Foundational Learning Assessments

4.1 Introduction

Over the past two decades, Africa has witnessed an unprecedented expansion in the collection of foundational learning data. National assessments have become more widespread, regional programmes have matured, and household-based assessments now provide internationally comparable evidence on children's foundational reading and numeracy skills across diverse educational contexts. As Chapter 3 demonstrated, access to these data has also improved substantially. Public-use microdata are now available for an increasing number of assessments, creating new opportunities for researchers, governments and development partners to undertake independent analyses of children's learning.

This progress has transformed what can be asked of assessment data.

Historically, most analyses relied on published reports, country summaries or a small set of headline indicators. These products remain essential for monitoring progress and communicating results, but they necessarily provide only a partial view of children's learning. Publicly available microdata allow researchers to move beyond these summaries by examining the underlying distributions, linking learning outcomes to children's household circumstances and investigating methodological questions that cannot be addressed using aggregated statistics alone.

The value of open microdata therefore extends well beyond transparency and reproducibility. They create opportunities to improve how assessment results are interpreted, compared and ultimately used to inform educational policy. Secondary analyses can identify sources of bias, evaluate assumptions underlying international comparisons, investigate patterns of inequality and generate evidence that contributes not only to our understanding of children's learning but also to improvements in assessment design itself.

This chapter illustrates these opportunities using two nationally representative household assessment programmes introduced in Chapter 3: UNICEF's Multiple Indicator Cluster Surveys (MICS) and the PAL Network's International Common Assessment of Numeracy and Reading (ICAN-ICAR). These programmes were selected because they combine nationally representative household samples with publicly available microdata and rich contextual information, making them particularly well suited to methodological investigation. The purpose of this chapter is not to compare the two programmes or to determine which provides the "better" assessment. Rather, they are used as complementary examples to demonstrate the kinds of questions that become possible when assessment microdata are openly available.

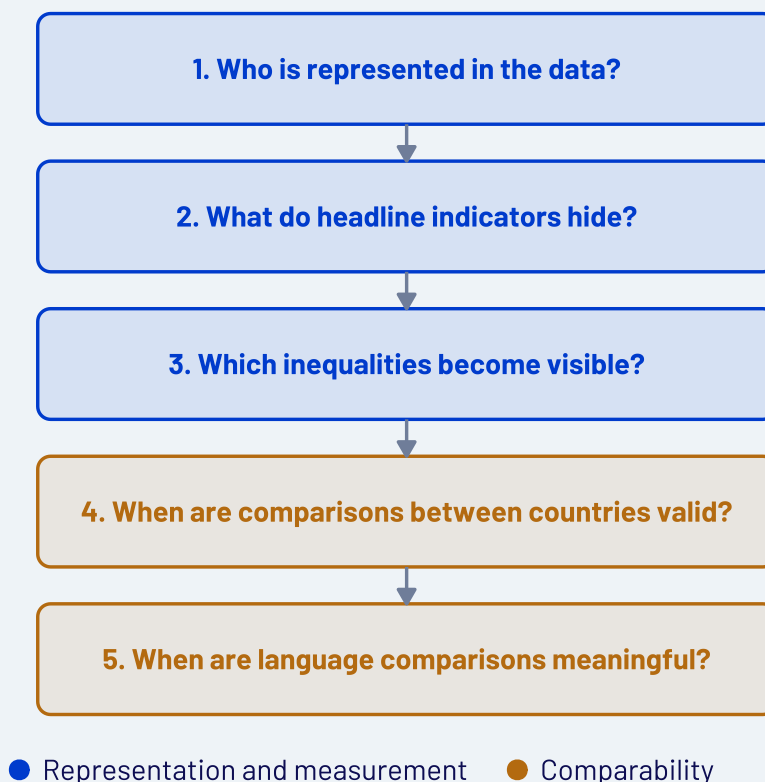
Throughout the chapter we move progressively from describing children's learning to examining the assumptions that underpin comparative analyses. Rather than asking simply whether learning differs across countries, we ask how researchers can determine whether those differences are meaningful and what additional evidence microdata provide to support their interpretation.

The chapter is organised around five methodological questions (Figure 4.1).

**FIGURE 4.1**

Open microdata allow us to move beyond headline indicators and interrogate how assessment results should be interpreted

Five questions for interpreting household assessment microdata



Together these questions illustrate a broader principle running throughout this report. Producing high-quality assessment data is only the first step in generating evidence about children's learning. Equally important is ensuring that those data remain accessible for rigorous secondary analysis. Publicly available microdata enable researchers not only to replicate published findings but also to investigate new questions, evaluate methodological assumptions and strengthen the interpretation of assessment results. In doing so, they improve the quality of the evidence available to policymakers while contributing to the continued development of foundational learning assessments themselves.

4.2 Household assessment microdata as a resource for comparative research

To illustrate the analytical opportunities created by publicly available household assessment microdata, this chapter draws on two complementary assessment programmes: UNICEF's Multiple Indicator Cluster Surveys (MICS) and the PAL Network's International Common Assessment of Numeracy and Reading (ICAN-ICAR). Although both are nationally representative household-based assessments, they were developed for different purposes and represent distinct approaches to measuring foundational learning.



The Multiple Indicator Cluster Surveys

MICS integrates a foundational learning assessment within a broad household survey designed primarily to monitor the wellbeing of women and children. Rather than functioning as a dedicated education assessment, it combines information on children's learning with detailed data on household characteristics, health, nutrition, education and living conditions. This integration allows learning outcomes to be analysed alongside one of the richest collections of socioeconomic information currently available for African countries.

The Foundational Learning Skills module assesses whether children have acquired the foundational reading and numeracy skills expected by the end of the early primary grades. Numeracy is assessed through four progressively more demanding components—reading numbers, number discrimination, simple addition and number patterns—while reading combines decoding and comprehension tasks appropriate to Grade 2 level. The module is administered to children aged 7–14 selected through nationally representative household sampling procedures.

A major strength of MICS is its commitment to open science. Public-use datasets, questionnaires, interviewer manuals, technical documentation and sampling reports are routinely released, allowing researchers to reproduce published findings, undertake secondary analyses and combine learning outcomes with the extensive contextual information collected by the survey.

MICS datasets and documentation are available through the UNICEF MICS survey portal, while harmonised multi-country datasets are available through the IPUMS-MICS project. Researchers should note, however, that in countries where multiple reading passages were administered in different languages, the IPUMS harmonised files retain only the first reading passage. These files are therefore not appropriate for analyses of reading achievement in multilingual countries, and researchers undertaking such analyses should instead use the original country-specific MICS datasets.

UNICEF is also increasingly demonstrating the analytical potential of the MICS Foundational Learning Skills data through a series of accessible, interactive analyses. These resources provide engaging examples of how the distinctive structure of MICS—combining direct assessment with information on children's schooling, households and wider circumstances—can be used to investigate the processes underlying observed learning outcomes, rather than reporting proficiency levels alone.¹

The International Common Assessment of Numeracy and Reading

ICAN-ICAR was developed by the PAL Network as a purpose-built assessment of foundational learning suitable for citizen-led and nationally representative household surveys. Unlike MICS, which embeds learning assessment within a broader household survey, ICAN-ICAR was designed specifically to measure children's foundational reading and numeracy using internationally comparable assessment frameworks.

The first nationally representative implementations were undertaken across 12 countries during 2025, including Kenya, Mali, Mozambique, Senegal, Tanzania and Uganda. Within sampled households, all children aged 5–16 were assessed individually by trained citizen assessors using standardised paper-based instruments. Alongside learning outcomes, the surveys collect extensive information on children's schooling, household assets, parental education and functional difficulties using the Washington Group Child Functioning Module.

The assessment framework was developed through extensive conceptual development, cognitive testing, field piloting and psychometric evaluation. Global proficiency thresholds were established using the Pairwise Comparison Method in collaboration with the Australian Council for Educational Research, while Item Response Theory scaling and differential item functioning analyses were used to ensure comparability across countries and population groups. The UNESCO Institute for Statistics has confirmed that the assessment satisfies the technical requirements for reporting Sustainable Development Goal Indicator 4.1.1(a).

In keeping with the PAL Network's commitment to transparency and knowledge sharing, the African ICAN-ICAR datasets are curated and disseminated through AFLEARN at DataFirst, University of Cape Town. Researchers have access to anonymised microdata together with standardised documentation, codebooks

1 UNICEF's interactive analyses are available through its dedicated [Foundational Learning Skills Resource page](#)



and metadata to support reproducible analyses and cross-country comparisons. By making these resources publicly available, the programme seeks to maximise the value of the data for research, policy and capacity building beyond the publication of national assessment reports.

Programme-specific proficiency benchmarks

Both assessment programmes classify children according to whether they have acquired the foundational skills expected by the end of the early primary grades. The methods used to establish these thresholds, however, differ.

The original MICS Foundational Learning Skills module was developed with the intention of supporting SDG 4.1.1(a) reporting. Within the assessment, children are classified as having acquired foundational numeracy only if they correctly answer all twenty-one mathematics items, while the reading benchmark combines successful decoding with literal and inferential comprehension. Although these thresholds reflected UNICEF's intended reporting framework, the UNESCO Institute for Statistics subsequently concluded that the original module did not fully satisfy the technical requirements for global SDG reporting. UNICEF has since developed the enhanced Foundational Learning Module (FLM 2.0), which aligns with the UIS reporting criteria and will be implemented in the seventh round of MICS.

ICAN-ICAR followed a different development pathway. Its proficiency thresholds were established explicitly for international reporting using psychometric standard-setting procedures and IRT-based scaling.

Because the two programmes use different assessment frameworks and different approaches to defining foundational learning, this chapter does not attempt to harmonise or equate their thresholds. Instead, each assessment is analysed using its own programme-specific proficiency benchmark. Throughout the remainder of the chapter, this term refers to the benchmark defined by each programme for identifying children who have acquired foundational reading or numeracy skills. Using this common terminology allows findings from the two programmes to be discussed consistently while recognising that the benchmarks themselves are not directly comparable.

The remainder of this chapter uses these two assessment programmes to illustrate how publicly available microdata can improve the interpretation, quality and comparability of foundational learning assessments.

Question 1. Who is represented in the data?

Every assessment tells the story of the population it measures. Before interpreting differences in learning outcomes, it is therefore essential to understand who is represented in the assessment—and equally importantly, who is not. This is one of the first questions that secondary analysts should ask when working with any assessment dataset.

Published assessment reports typically present learning results for the population included in the assessment but provide relatively little information about children who fall outside the target population or who do not ultimately contribute to the reported learning indicators. Publicly available microdata make it possible to investigate these questions directly, allowing researchers to examine eligibility, participation, response patterns and sample composition in much greater detail.

Household-based assessments are particularly valuable because they begin with a representative sample of households rather than schools. They therefore include children who are currently enrolled, those who have dropped out and those who have never attended school. In addition, they collect detailed information on children's schooling histories, school types and participation in different assessment components. Together, these data provide a much richer understanding of representativeness than is possible from published reports alone.

This section examines four dimensions of representation that can be investigated using household assessment microdata:

- Who participates in school?
- What types of schools do children attend?



- Who ultimately completes the assessment?
- How do realised samples influence interpretation?

Together these examples demonstrate that representativeness is not established solely through sample design. It is the outcome of a sequence of decisions relating to school participation, eligibility, assessment design and response patterns, each of which influences who ultimately contributes to the reported learning statistics.

Response rates and realised samples

The first dimension of representativeness concerns the distinction between the intended sample and the realised analytical sample. A representative sampling design does not automatically guarantee a representative analytical sample. Non-response at any stage of data collection can alter the composition of the realised sample and therefore influence the interpretation of assessment results.

Household surveys typically involve multiple stages of participation, including completion of the household interview, identification of eligible children and successful completion of the learning assessment itself. Because participation may differ across countries and across assessment components, analysts should examine realised response rates before drawing substantive conclusions.

Perhaps surprisingly, nationally representative household-based learning assessments achieve response rates that are comparable to, and in many cases exceed, those reported for large-scale school-based assessments.

Table 4.1 presents household, and child assessment response rates for the MICS surveys. Household response rates exceed 90% in the majority of participating countries. For the mathematics assessment, final child response rates range from 68% in Lesotho to 98% in Guinea-Bissau, with half of the participating countries achieving response rates above 90%². This demonstrates that nationally representative household learning assessments can consistently achieve exceptionally high participation. Although overall response rates for numeracy are generally high, countries with lower participation warrant additional investigation. Publicly available microdata make it possible to examine whether non-response is systematically associated with characteristics such as age, sex, location or household wealth, allowing researchers to evaluate the potential for non-response bias rather than relying solely on overall response rates.

Reading presents a different picture.

Across almost all participating countries, realised response rates for the reading assessment are substantially lower than for numeracy. Importantly, this difference arises not primarily because children refuse to participate in the survey, but because many children are either ineligible to attempt the reading assessment owing to language mismatch or choose not to attempt the reading passage. Consequently, the realised reading sample differs systematically from both the intended survey population and the realised mathematics sample.

In The Gambia, for example, approximately 19% of Grade 2 children are excluded because no suitable assessment language is available, while a further 10% decline to attempt the reading passage despite completing the mathematics assessment. Similar patterns occur in several other countries.

These different mechanisms of non-response have important implications for interpreting assessment results. A comparison of mathematics performance among children who did and did not complete the reading assessment illustrates the extent of this selection. Children who completed the reading assessment achieved an average mathematics score of 66%, compared with only 22% among those who did not—a difference of 43 percentage points. Similarly, 18% of children completing the reading assessment reached the programme-specific proficiency benchmark for mathematics compared with just 1% among those who did not.

These findings demonstrate that children completing the reading assessment constitute a highly selected subgroup of the overall survey sample. Researchers should therefore exercise caution when comparing reading and mathematics outcomes within countries because the two assessments represent different realised populations.

² Tunisia is the one country with two MICS6 surveys. Due to the substantially lower response rates in the 2023 survey, the analyses in the remainder of this chapter use the 2018 survey for Tunisia. The only exception is the analysis on language of instruction as the 2018 survey did not include the relevant questions.

**TABLE 4.1.****Household and child assessment response rates in MICS**

Country	Year	Household response rate	Final child-level response rate	
			Mathematics assessment	Reading assessment
Benin	2021	96%	93%	73%
CAR	2018	90%	81%	45%
Chad	2019	99%	95%	31%
Comoros	2022	91%	81%	68%
DRC	2017	100%	98%	66%
Eswatini	2021	89%	81%	76%
Gambia	2018	96%	88%	54%
Ghana	2017	98%	95%	78%
Guinea-Bissau	2018	98%	97%	35%
Lesotho	2018	85%	68%	63%
Madagascar	2018	89%	74%	63%
Malawi	2019	95%	87%	63%
Nigeria	2021	95%	91%	69%
Sao Tome & Principe	2019	92%	82%	71%
Sierra Leone	2017	98%	94%	43%
Togo	2017	94%	88%	79%
Tunisia	2018	94%	88%	87%
Tunisia	2023	81%	66%	64%
Zimbabwe	2019	92%	87%	82%

Notes: The final child response rate is calculated as the product of the household response rate, the response rate for the questionnaire administered to children aged 5–17 years, and the proportion of eligible children aged 7–14 years who completed the mathematics or reading assessment.

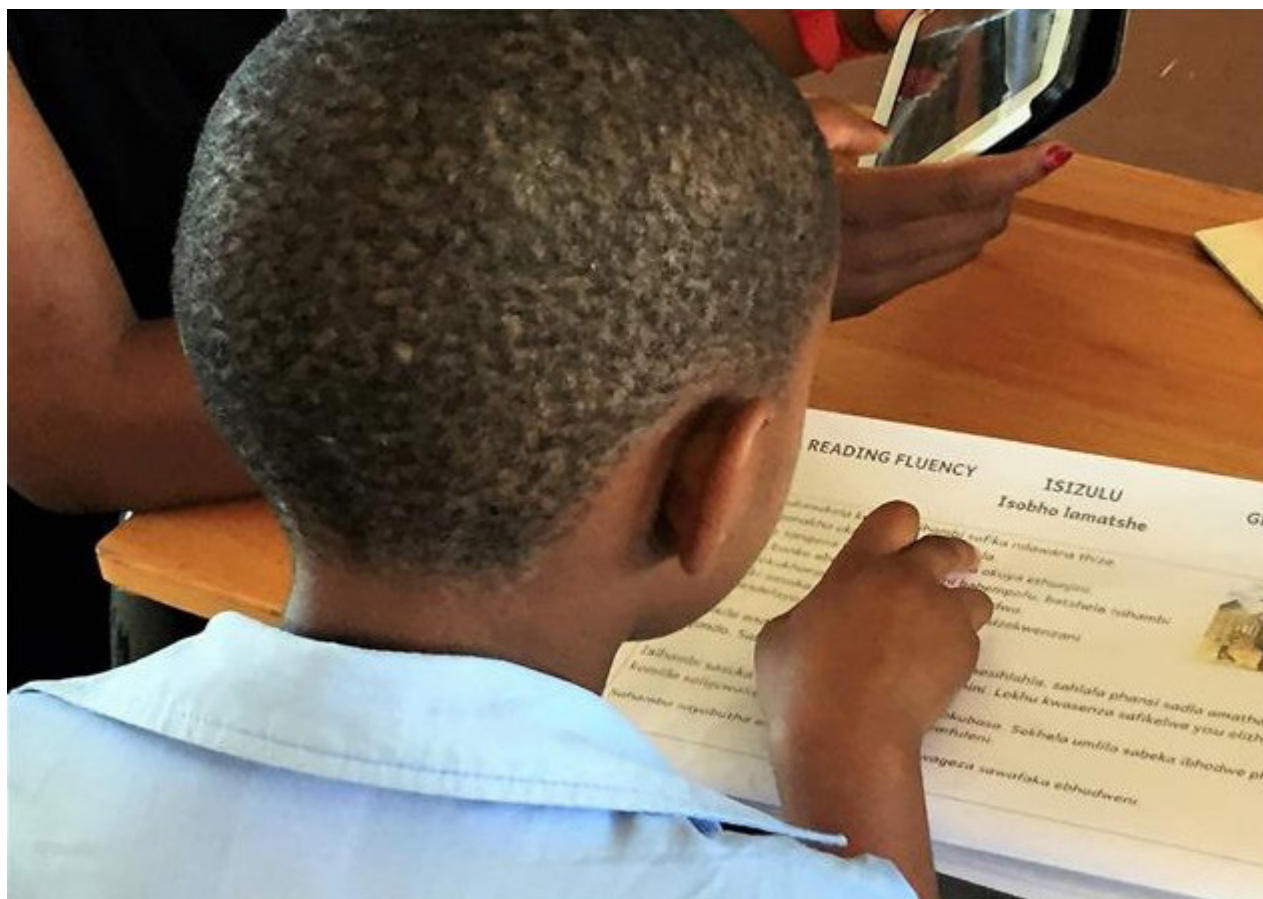
TABLE 4.2.**Household, child and assessment response rates in ICAN-ICAR**

	Household Response Rate (%)	Child is available to be assessed	Overall child-level response rate
Kenya	99.7%	94.2%	94.0%
Mali	97.9%	95.0%	93.0%
Mozambique	99.6%	99.1%	98.7%
Senegal	99.4%	88.8%	88.3%
Tanzania	99.6%	89.0%	88.6%
Uganda	99.8%	93.2%	93.0%

Notes: The overall child-level response rate is calculated as the product of the household response rate and the proportion of children who are available to be assessed.

Table 4.2 presents the corresponding response rates for the six African ICAN-ICAR countries. Overall child response rates are consistently high ranging from 88% in Senegal to 99% in Mozambique. However, household listing was undertaken separately from assessment in many countries, meaning that refusals during household listing are not captured in the reported response rates. As with any survey, analysts should therefore understand precisely how response rates have been calculated before comparing them across assessment programmes. Out-of-school children

The most obvious difference between household- and school-based assessments concerns who is eligible to



be assessed. School-based assessments necessarily describe children attending school, whereas household-based assessments begin with all children living in sampled households, regardless of whether they are enrolled. This seemingly simple distinction has profound implications for how assessment results should be interpreted.

Figure 4.2 presents out of school rates at ages 7, 10 and 16 across the African countries included in MICS and ICAN-ICAR.

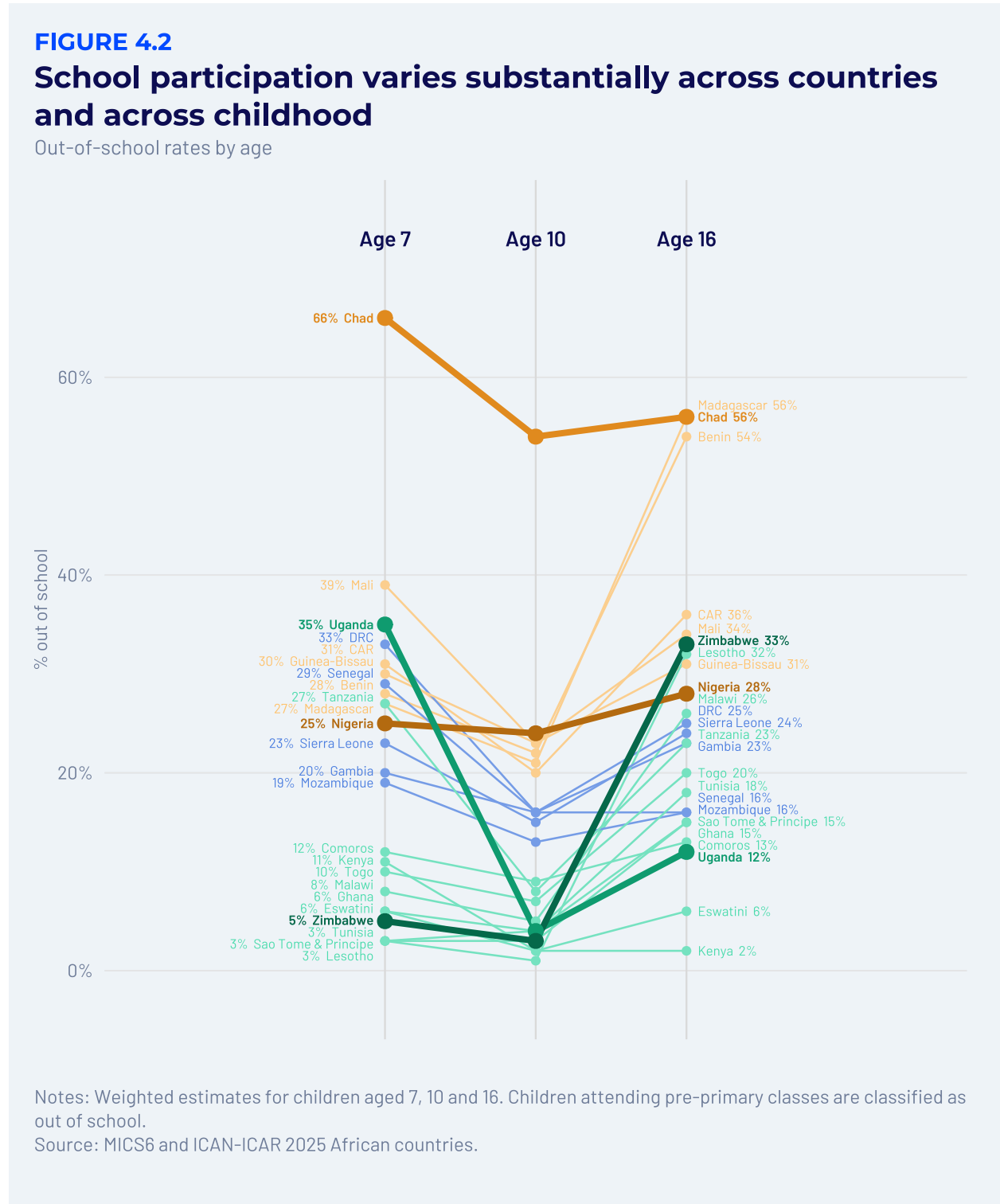
At age seven, there are substantial differences between countries in school participation, primarily reflecting variation in the age at which children enter school. By age ten, these differences instead reflect the extent to which countries have achieved universal participation in primary education. During adolescence, participation begins to diverge again as school dropout becomes increasingly common. The figure therefore illustrates not only how school participation changes across childhood but also how the representativeness of school-based assessments changes over the life course.

Between the ages of 10 and 16, the proportion of children who are out of school increases in almost every country, although the rate of increase varies considerably. Kenya and Senegal experience relatively little change over this period, indicating that most children who enter school remain enrolled throughout the primary years. By contrast, Zimbabwe, Lesotho, Benin and Madagascar exhibit particularly steep increases in school exclusion during adolescence. Nigeria and Comoros display comparatively flat participation profiles



across childhood, although this reflects persistently high rates of non-enrolment rather than sustained universal participation in Nigeria.

These patterns have important implications for interpreting school-based assessments. In countries where



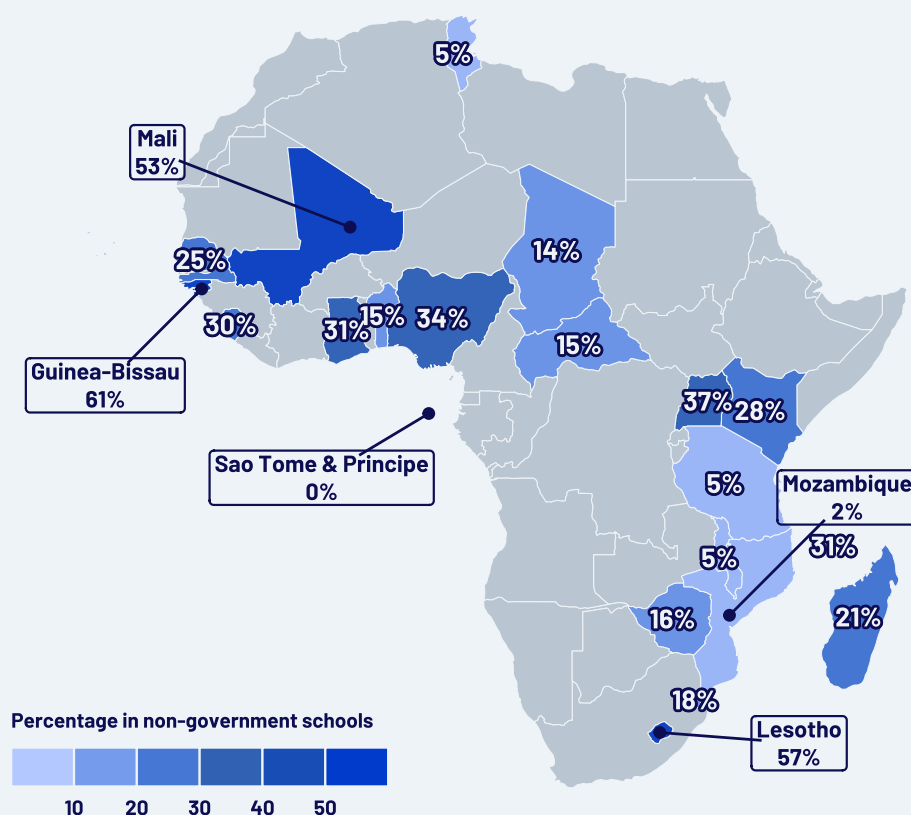
substantial numbers of children remain out of school, school-based assessments inevitably provide an incomplete picture of national learning because they exclude some of the most educationally disadvantaged children. Moreover, the extent of this exclusion differs markedly across countries and changes systematically with age, complicating direct comparisons between education systems.

Representation, however, depends not only on enrolment but also on attendance. Even among children who are enrolled, absenteeism is widespread across many African education systems. Evans and Acosta (2023), for example, document high levels of school absenteeism across the continent, implying that school-based assessments may fail to capture a substantial proportion of enrolled learners on the day of testing. Household-based assessments provide an important complement by capturing all children living in sampled households, regardless of their current school participation or attendance.

FIGURE 4.3

The role of non-government schools differs markedly across African education systems

Percentage of children attending non-government schools in grades 1 to 3



Notes: Weighted estimates for children enrolled in Grades 1–3. The figure excludes the Democratic Republic of Congo, The Gambia and Togo, as the MICS6 questionnaire in those countries did not include the question on school type.

Source: MICS6 and ICAN-ICAR 2025 African countries.

Representation beyond government schools

A third dimension of representativeness concerns the types of schools included in an assessment. Household surveys provide a complete picture of the school sectors attended by children, allowing researchers to quantify the contribution of government and non-government schools to national enrolment. This is important because school-based assessments may differ in the extent to which they include different school sectors



within their sampling frames. Consequently, assessments that appear to target the same population—for example, Grade 2 learners—may nevertheless represent different segments of the education system.

Household-based assessments therefore provide an independent benchmark against which the representativeness of school-based assessment samples can be evaluated. Understanding the composition of the assessed population is another essential step in interpreting assessment results and making valid comparisons across countries.

Figure 4.3 shows the proportion of children attending non-government schools across grades in the participating countries. The importance of the non-government sector varies enormously across Africa. More than half of children enrolled in the early grades attend non-government schools in Guinea-Bissau (60%), Lesotho (57%) and Mali (53%), while fewer than 5% do so in Mozambique, Tunisia, Malawi and São Tomé and Príncipe.

In most countries, the share of children attending non-government schools remains relatively stable across grades. There are, however, several notable exceptions. In Madagascar and Comoros, the proportion attending non-government schools increases substantially, rising from 19% and 28% respectively in Grade 1 to 51% and 64% by Grade 9. Kenya exhibits the opposite pattern, with the proportion attending non-government schools declining from 33% in Grade 1 to just 12% by Grade 9. These contrasting trajectories suggest important differences in the structure of education systems and the roles played by government and non-government providers over children's school careers.

Household survey data can also be compared directly with school-based assessments to examine whether different assessment programmes represent similar school sectors. Table 4.3 compares the proportion of Grade 2 children attending non-government schools in the household-based MICS and ICAN-ICAR surveys with estimates from the school-based PASEC (Grade 2) and AMPL (Grade 3) assessments.

TABLE 4.3.

Share of Grade 2 children attending non-government schools: comparison between household-based and school-based assessments

	Household-based		School-based	
	% in non-government schools	Source	% in non-government schools	Source
Benin	15%	MICS 2021	28%	PASEC 2019
Madagascar	22%	MICS 2018	63%	PASEC 2019
Senegal	27%	ICAN-ICAR 2025	15%	PASEC 2019
Chad	14%	MICS 2019	26%	PASEC 2019

Notes: Weighted estimates for children enrolled in grade 2.

The comparisons reveal substantial differences between assessment programmes. In Senegal, for example, 27% of Grade 2 children in ICAN-ICAR attend non-government schools compared with only 15% in the PASEC 2019 Grade 2 sample. In Benin, the opposite pattern is observed, with private schooling more common in the PASEC sample (28%) than in MICS (15%).

The reasons for these differences require further investigation and may reflect differences in sampling frames, definitions of non-government schools or changes in school sector composition over time. More broadly, they demonstrate that representativeness extends beyond the characteristics of individual children. It also depends on whether the institutions in which children learn are represented in comparable ways across assessment programmes.



WHY MICRODATA MATTER

Most published assessment reports describe the intended sample but provide relatively limited information about the population ultimately represented in the reported learning outcomes. Publicly available household assessment microdata allow researchers to examine every stage of representation—from school participation and school sector to assessment eligibility and realised response rates. This transforms representativeness from an assumption into an empirical question, strengthening both the interpretation of learning outcomes and the validity of comparisons across countries.

Question 2. What do headline indicators hide?

The previous section demonstrated that understanding who is represented in an assessment is a prerequisite for interpreting learning outcomes. Once the analytical population has been established, the next question is how learning itself should be summarised.

Most assessment reports rely on a relatively small number of headline indicators, typically average scores and the proportion of children reaching a defined proficiency benchmark. These measures play an essential role in monitoring progress, comparing countries and communicating results to policymakers. Their strength lies in their simplicity. A single statistic can summarise the performance of thousands of children and provide an accessible way of tracking progress over time or comparing education systems.

Every summary statistic, however, conceals information.

Average scores compress an entire distribution of achievement into a single number, while proficiency benchmarks reduce a continuum of learning to a binary classification. Countries that appear similar according to headline indicators may therefore differ substantially in the underlying distribution of learning or in the specific competencies that children have mastered.

Publicly available microdata allow researchers to move beyond these summaries. Rather than asking simply how many children reach the programme-specific proficiency benchmark, analysts can examine where children lie along the learning continuum and which skills they have acquired. This provides a much richer understanding of children's learning while also improving the interpretation of commonly reported indicators.

This section illustrates two complementary examples. First, it examines how learning distributions reveal important patterns hidden by averages and proficiency benchmarks. Second, it demonstrates how analysing individual assessment components helps identify the specific competencies that children have—and have not—yet mastered.

Together these analyses show that headline indicators should be viewed as the beginning of analysis rather than its end.

Looking beyond average scores and proficiency benchmarks

Programme-specific proficiency benchmarks have become one of the most widely reported indicators in international education monitoring because they provide a common reference point for tracking progress towards Sustainable Development Goal 4.1.1. Their appeal lies in their simplicity. Children either reach the benchmark or they do not.

This simplicity also creates important limitations.

The indicator does not distinguish between children who narrowly miss the benchmark and those who remain at the earliest stages of skill acquisition. This limitation becomes particularly important in countries where relatively few children reach the benchmark. In such settings, reporting only the proportion meeting the benchmark provides relatively little information about the substantial variation that exists below it.

Publicly available microdata make it possible to examine the entire learning continuum rather than a single point upon it.



Figure 4.4 presents the distribution of total mathematics scores (out of 21) for children enrolled in Grades 2 and 3 across the eighteen African MICS countries. The shaded red portion of each distribution represents children who reached the programme-specific proficiency benchmark. Countries are ordered according to the proportion of children meeting this benchmark, from the Democratic Republic of the Congo, where only 0.2% of children achieved it, to São Tomé and Príncipe, where approximately 18% did so.

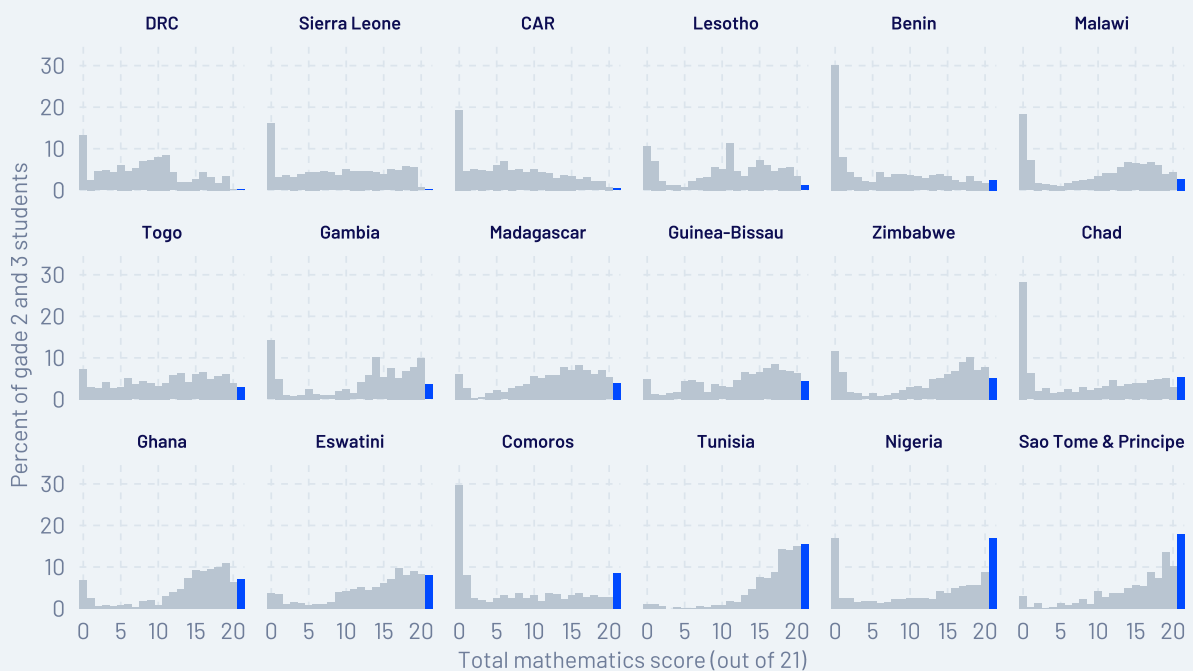
Several observations emerge.

FIGURE 4.4

Similar headline indicators can conceal very different learning distributions

Distribution of mathematics scores for children enrolled in grades 2 and 3

Blue bars show children who reach the MICS proficiency benchmark



Notes: Weighted numeracy score distributions for children enrolled in Grades 2 and 3. Countries are ordered by the proportion reaching the benchmark.

Source: MICS6 African countries.

First, the benchmark captures only a small proportion of the learning distribution in every country. Although it represents the level of numeracy expected by the end of the early primary grades, the vast majority of children enrolled in Grades 2 and 3 remain below it. Focusing exclusively on whether children meet the benchmark therefore tells us relatively little about the learning levels of most children.

Second, countries with similar proficiency rates often exhibit very different distributions below the benchmark.

Tunisia and Nigeria provide a particularly striking example. In both countries approximately 16–17% of children enrolled in Grades 2 and 3 reach the programme-specific proficiency benchmark. Yet the distributions beneath this threshold differ markedly. In Nigeria, a substantial proportion of children score zero on the assessment, whereas very few children do so in Tunisia. Instead, the Tunisian distribution is concentrated immediately below the benchmark.

These contrasting distributions imply very different educational challenges. Although the two countries appear almost identical when viewed through the lens of the programme-specific proficiency benchmark,



Tunisia faces the task of helping many children make the final step towards proficiency, whereas Nigeria must address the needs of a much larger group of children who have yet to acquire even the most basic numeracy skills. Similar proficiency rates therefore do not necessarily imply similar policy priorities.

Third, average scores can conceal equally important differences.

Comoros and the Democratic Republic of the Congo, for example, have almost identical average mathematics scores of around 8 out of 21. Yet these similar national averages mask strikingly different learning distributions. Comoros has the fourth highest proportion of children reaching the programme-specific proficiency benchmark (around 9%), but almost one-third of children score zero on the assessment—a level matched only by Benin. In contrast, virtually no children in the Democratic Republic of the Congo reach the benchmark, yet achievement is distributed much more evenly across the lower part of the score distribution.

These examples illustrate that programme-specific proficiency benchmarks and average scores each provide only one perspective on children's learning. Neither provides a complete description. Access to microdata allows researchers to examine the full distribution of achievement, providing richer evidence about learning while offering policymakers more informative guidance about where interventions may be most needed.

Looking inside the assessment

The previous analysis treated learning as though it could be represented by a single score. In practice, however, foundational reading and numeracy develop through the progressive acquisition of multiple competencies. One of the greatest strengths of publicly available microdata is that they allow analysts to move beyond overall scores and examine performance on individual assessment components.

This is particularly valuable in foundational learning because both reading and numeracy consist of multiple related competencies that develop progressively over time. Numeracy assessments typically move from recognising numbers and comparing quantities to performing arithmetic and identifying numerical patterns. Reading assessments similarly progress from oral language and letter-sound knowledge to decoding, word reading and comprehension.

Overall scores provide little insight into where children begin to struggle. Microdata make it possible to identify these transition points.

Such analyses can inform curriculum development, teacher support and instructional improvement while also contributing to assessment design by identifying tasks that discriminate effectively across different levels of ability.

Mathematics competencies

Figure 4.5 presents performance among children enrolled in Grades 2 and 3 on the four mathematics components included in the MICS Foundational Learning Skills module. For each component, the red bar represents children who were unable to answer any items correctly, while the green bar represents those who completed every item successfully. All other children are represented by the grey area. Countries are ordered according to the proportion of children scoring zero on each component.

Viewing the assessment in this way begins to unpack where children fall behind and what they can—and cannot—yet do. It also presents learning in a way that is often more intuitive for teachers, curriculum specialists and policymakers than overall assessment scores.

Again, Tunisia and Nigeria provide an informative comparison.

For the first component, reading numbers, 19% of Nigerian children enrolled in Grades 2 and 3 were unable to recognise even the simplest number, whereas almost all children in Tunisia recognised at least some numbers and 76% successfully completed all six items.

The contrast continues in number discrimination, where children identify the larger of two numbers. Around 63% of Tunisian children completed every item successfully compared with only 36% in Nigeria.



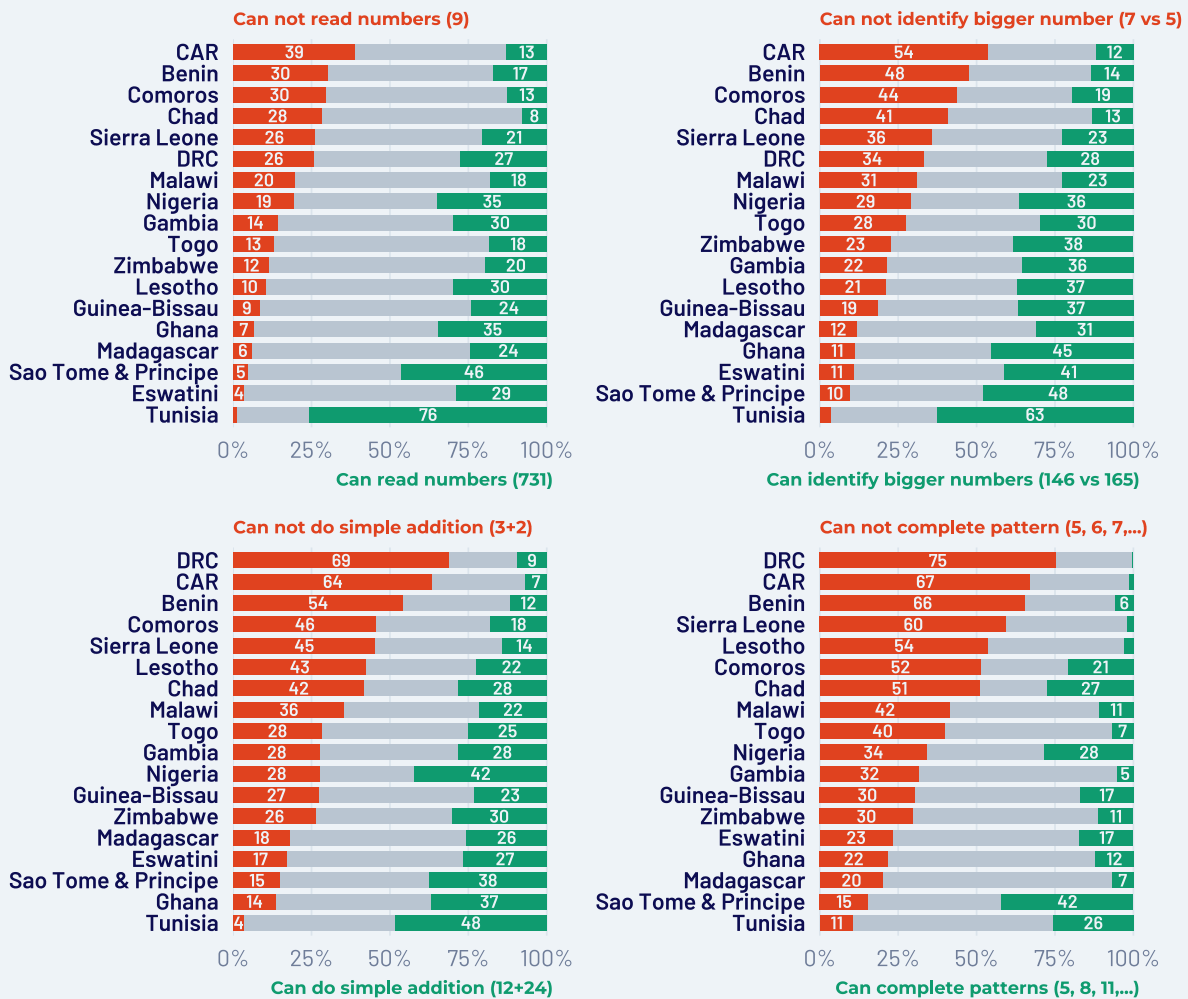
Performance becomes more similar for simple addition, where approximately 42–48% of children in both countries completed all five items correctly. By the final component, number patterns, similar proportions of children in the two countries completed every item, explaining why the overall programme-specific proficiency rates appear almost identical. Yet an important difference remains: 34% of Nigerian children were unable to complete even the simplest pattern compared with only 11% in Tunisia.

FIGURE 4.5

Looking inside the assessment reveals where children's numeracy begins to break down

Performance by mathematics competency for children enrolled in grades 2 and 3

None correct · Some correct · All correct



Notes: Weighted estimates for children enrolled in Grades 2 and 3. For each assessment component, the figure distinguishes children answering no items correctly, all items correctly, and those falling between these extremes. Countries are ordered by the proportion scoring zero on each component.

Source: MICS6 African countries.

The progression across components is also evident within individual countries. In Benin, for example, approximately 30% of children were unable to read numbers, 48% could not identify the larger of two numbers, 54% could not complete simple addition and fully 66% could not solve the simplest number pattern.

Because children must complete every mathematics component successfully to reach the programme-



specific proficiency benchmark, relatively few children achieve this standard. The component-level analysis, however, reveals where the majority of children are situated along the learning pathway and therefore provides considerably more information than the benchmark alone.

Reading competencies

The ICAN-ICAR data allow a similar analysis for foundational reading.

Although relatively few children enrolled in Grades 2 and 3 reach the programme-specific reading proficiency benchmark, performance varies substantially across the underlying reading competencies.

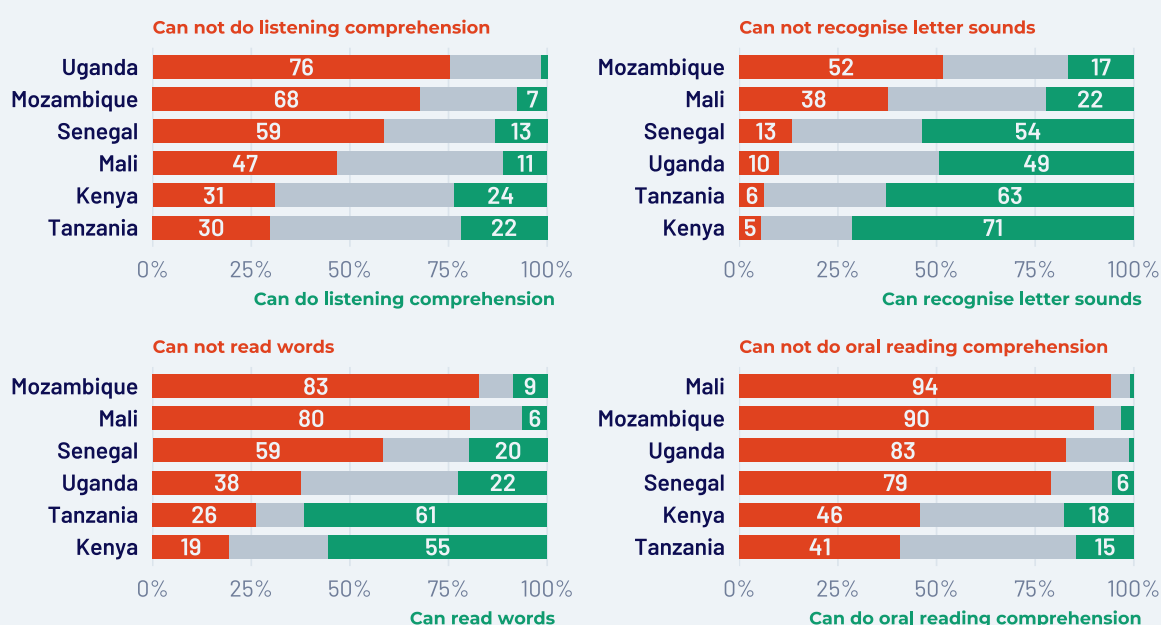
Figure 4.6 presents the proportion of children scoring zero and 100% on four reading components: listening comprehension, letter-sound recognition, word reading and oral reading comprehension.

FIGURE 4.6

Reading profiles reveal very different points at which children's learning breaks down

Performance by reading competency for children enrolled in grades 2 and 3

None correct · Some correct · All correct



Notes: Weighted estimates for children enrolled in Grades 2 and 3. For each assessment component, the figure distinguishes children answering no items correctly, all items correctly, and those falling between these extremes. Countries are ordered by the proportion scoring zero on each component.

Source: ICAN-ICAR 2025 African countries.

These profiles reveal where children's reading development begins to break down.

Uganda provides a particularly informative example. Oral language skills appear especially underdeveloped, with 76% of children scoring zero on the listening comprehension component. Performance improves substantially for basic decoding, with around half of children recognising every letter sound and approximately 62% able to read at least one word. Performance deteriorates sharply, however, once children are required to understand connected text. Fully 83% score zero on oral reading comprehension.



This pattern is consistent with the cumulative nature of reading development. Reading comprehension depends not only on decoding but also on vocabulary, oral language and language comprehension.

In Uganda, where the assessment is conducted in English, these findings suggest that limited familiarity with the language of assessment may constitute an additional barrier to demonstrating reading proficiency. This issue is explored further in Question 5.

More broadly, these analyses illustrate that overall assessment scores often conceal important information about where learning begins to falter. Examining individual competencies allows researchers to identify specific stages in children's learning progression where instructional support may be most needed.

WHY MICRODATA MATTER

Published assessment reports necessarily summarise learning using a small number of headline indicators. Publicly available microdata allow researchers to move beyond these summaries by examining both the distribution of learning and the individual competencies that underpin overall scores. This transforms assessments from tools for monitoring performance into resources for understanding how children learn, where learning begins to break down and which educational challenges require the greatest attention.

Question 3. Which inequalities become visible?

The previous section showed that publicly available microdata allow researchers to move beyond headline indicators by examining the full distribution of learning and the specific competencies children have acquired. An equally important advantage of household assessment microdata is that they allow learning outcomes to be linked directly to children's household circumstances.

Published assessment reports typically present only a limited number of subgroup comparisons, often disaggregating results by sex or place of residence. Household assessment datasets contain much richer contextual information, including household wealth, parental education, school participation and children's educational histories. Linking these variables to learning outcomes makes it possible to investigate not only whether inequalities exist, but also how they arise and how their interpretation depends on the outcome measure being used.

This section illustrates two examples.

First, it examines how wealth inequalities appear very different depending on whether learning is summarised using programme-specific proficiency benchmarks or average scores. Second, it demonstrates how household assessments uniquely allow researchers to investigate the relationship between school participation and learning by comparing children who are enrolled, have left school and have never attended school.

Together these examples illustrate that publicly available microdata deepen our understanding of educational inequality while revealing dimensions of disadvantage that remain largely invisible in published assessment reports.

Wealth-related inequalities

Household wealth remains one of the strongest predictors of educational outcomes across Africa. Because both MICS and ICAN-ICAR collect detailed information on household assets and living conditions, learning outcomes can be examined across nationally comparable wealth groups. These analyses reveal substantial socioeconomic gradients in both reading and numeracy. They also demonstrate that the interpretation of inequality depends critically on how learning is measured.

Where overall performance is very low, differences in the proportion of children reaching the programme-specific proficiency benchmark may underestimate the true extent of socioeconomic disparities because few

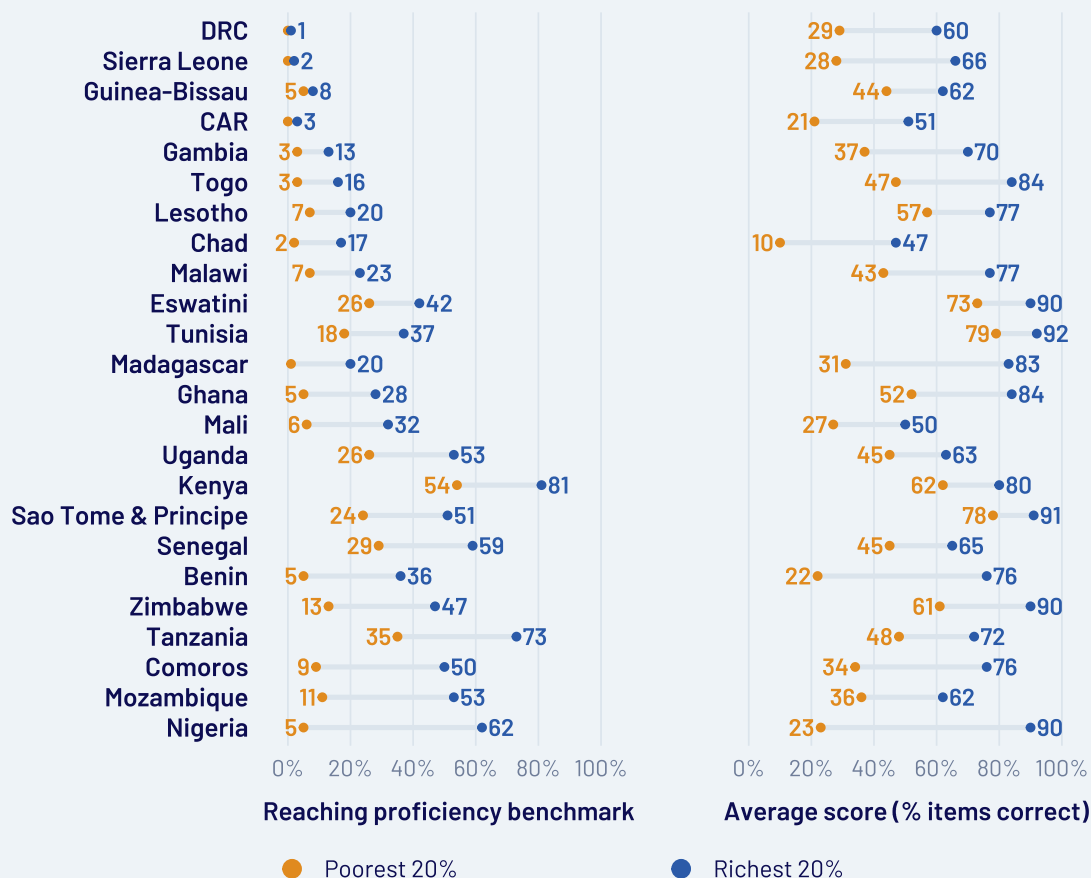


children from any wealth group achieve the threshold. Average scores and score distributions often reveal much larger inequalities.

FIGURE 4.7

Proficiency benchmarks can mask substantial wealth inequalities in learning

Wealth inequalities in mathematics: programme-specific proficiency benchmarks and average scores for children aged 7 to 14



Notes: Weighted estimates for children aged 7–14 living in the poorest and richest 20% of households. The left panel shows the proportion reaching each programme's proficiency benchmark; the right panel shows average mathematics scores as the percentage of assessment items answered correctly. Countries are ordered by the size of the wealth gap in reaching the proficiency benchmark. For MICS, wealth is measured using UNICEF's standard household wealth index. For ICAN-ICAR, an equivalent wealth index was constructed using principal components analysis of household assets and access to basic services. Classical test scores are used for ICAN-ICAR in the right panel to place both assessments on a percentage-correct scale; MICS and ICAN-ICAR achievement levels should not be compared directly. Source: MICS6 and ICAN-ICAR 2025 African countries.

Figure 4.7 compares mathematics performance among children from the poorest and richest 20% of households across the African MICS and ICAN-ICAR countries. The left-hand panel presents the proportion of children reaching each programme's proficiency benchmark, while the right-hand panel presents average mathematics scores expressed as the percentage of assessment items answered correctly.

The purpose of the figure is not to compare achievement between MICS and ICAN-ICAR, whose scales are



not directly comparable, but rather to compare wealth inequalities within each country. For MICS, household wealth is measured using UNICEF's standard wealth index. For ICAN-ICAR, a comparable wealth index was constructed separately for each country using principal components analysis based on household assets and access to services before households were classified into wealth deciles. Although ICAN-ICAR also provides Item Response Theory proficiency scores, classical test scores are used here to place both assessments on a common percentage-correct scale. Analyses using ICAN-ICAR alone should generally rely on the IRT proficiency scores.

Several findings emerge.

First, wealth inequalities are substantial across almost every participating country regardless of how learning is measured.

Second, the choice of outcome measure matters.

The left-hand panel suggests relatively modest wealth differences in several countries because very few children—whether rich or poor—reach the programme-specific proficiency benchmark. In the Democratic Republic of the Congo and Sierra Leone, for example, virtually no children from the poorest households achieve the benchmark, while only around one percent of those from the richest households do so.

The right-hand panel tells a very different story. Although few children from either wealth group reach the benchmark, children from the richest households consistently score much higher across the assessment. The benchmark therefore masks substantial differences in learning that become visible only when average scores - or, ideally, the full score distribution- are examined.

The magnitude of wealth disparities is striking. In Nigeria, children from the richest households achieve an average mathematics score of 90%, compared with only 23% among children from the poorest households. The smallest wealth gaps are observed in Eswatini, São Tomé and Príncipe and Tunisia, although even here children from the richest households outperform those from the poorest by around 13 percentage points.

The broader methodological lesson is that different outcome measures answer different questions. Programme-specific proficiency benchmarks identify whether children have reached an agreed minimum level of competence, whereas average scores describe where children lie along the learning continuum. Examining both provides a much richer understanding of educational inequality than either measure alone.

Learning and school participation

Household surveys provide another unique opportunity for understanding educational inequality by making visible children who are excluded from school.

School-based assessments necessarily describe only children who are present in school at the time of testing. Household assessments, by contrast, include children who have never attended school, those who previously attended but have dropped out and those who are currently enrolled. This allows learning to be interpreted alongside children's educational pathways rather than independently of them.

The first step is to examine how school participation itself varies across socioeconomic groups.

Figure 4.8 compares school enrolment among children aged 7 to 14 from the poorest and richest 20% of households.

Two patterns stand out immediately.

First, enrolment among children from the richest households is consistently high across almost all participating countries. In contrast, enrolment among children from the poorest households varies enormously, ranging from almost universal participation in some countries to fewer than one-third of children enrolled in others.

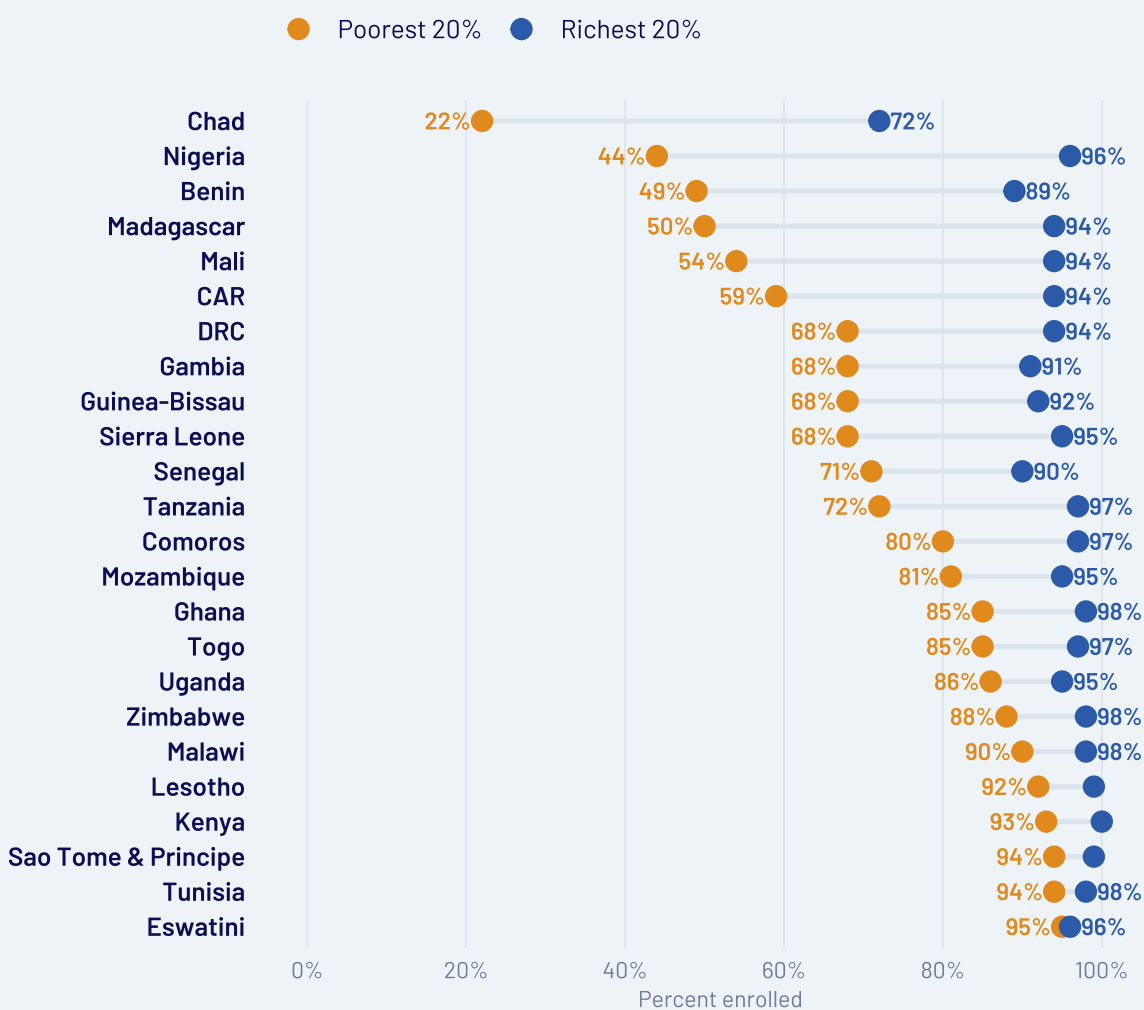
Second, wealth gaps in school participation are substantial in many education systems. In Nigeria, for example, only 41% of children from the poorest households are enrolled compared with 95% of children from the richest households.

These findings suggest that some of the wealth inequalities observed in learning outcomes may themselves reflect inequalities in school participation. Household assessment microdata allow this relationship to be investigated directly.

FIGURE 4.8

Wealth gaps in school participation remain substantial in many countries

School enrolment by household wealth for children aged 7 to 14



Notes: Weighted estimates for children aged 7 to 14 living in the poorest and richest 20% of households. Countries are ordered by the percentage of children in the poorest 20% of households who are enrolled. For MICS, wealth is measured using UNICEF's standard household wealth index. For ICAN-ICAR, an equivalent wealth index was constructed using principal components analysis of household assets and access to basic services. Source: MICS6 and ICAN-ICAR 2025 African countries.

Before comparing the learning outcomes of enrolled and out-of-school children, it is important to account for the confounding effect of age. Kenya and Zimbabwe provide useful examples because they have very different age profiles of school participation. In Kenya, 11% of seven-year-olds are enrolled, but enrolment rises rapidly to almost universal levels by age 10 and remains high through age 16 (see Figure 4.2). As a result, the out-of-school population is disproportionately younger than the enrolled population. Zimbabwe shows the opposite pattern: 95% of seven-year-olds are enrolled, but participation declines steadily during adolescence, with



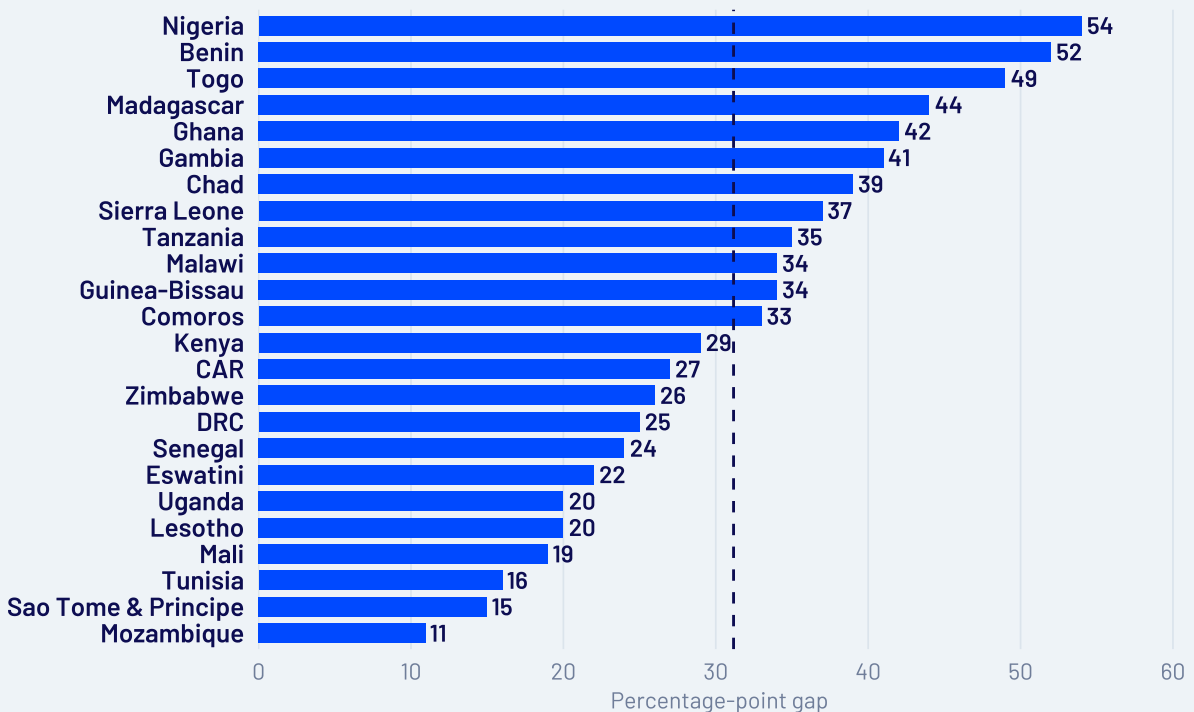
one-third of 16-year-olds out of school. Consequently, out-of-school children in Zimbabwe are, on average, older than those who remain enrolled. These contrasting age profiles substantially affect simple comparisons of learning outcomes. In Kenya, the unadjusted gap in average numeracy scores between enrolled and out-of-school children is 40 percentage points, but after controlling for age the estimated gap falls to 29 percentage points. In Zimbabwe, the reverse occurs: the unadjusted gap is 16 percentage points but increases to 26 percentage points once age differences are taken into account. This example illustrates a broader advantage of publicly available microdata. Rather than relying on descriptive comparisons, researchers can control for important confounding factors and produce estimates that more accurately reflect the relationship of interest.

Figure 4.9 shows, for each country the percentage point gap in mathematics scores between children who are currently enrolled with those who are not.

FIGURE 4.9

Out-of-school children have substantially lower mathematics scores in every country

Mathematics performance gap between enrolled and out-of-school children for children aged 7 to 14



Notes: Weighted estimated gap in mathematics scores, in percentage points, between enrolled and non-enrolled children aged 7–14. Estimates control flexibly for age using indicator variables for each year of age.

Source: MICS6 and ICAN-ICAR 2025 African countries.

Across every participating country, children attending school substantially outperform those who are out of school. These gaps range from 11 percentage points in Mozambique to just over 50 percentage points in Benin and Nigeria.

The category "out of school", however, conceals considerable diversity. It combines two distinct groups: children who have never attended school and children who previously attended but subsequently left. These groups often have very different educational experiences and learning trajectories.

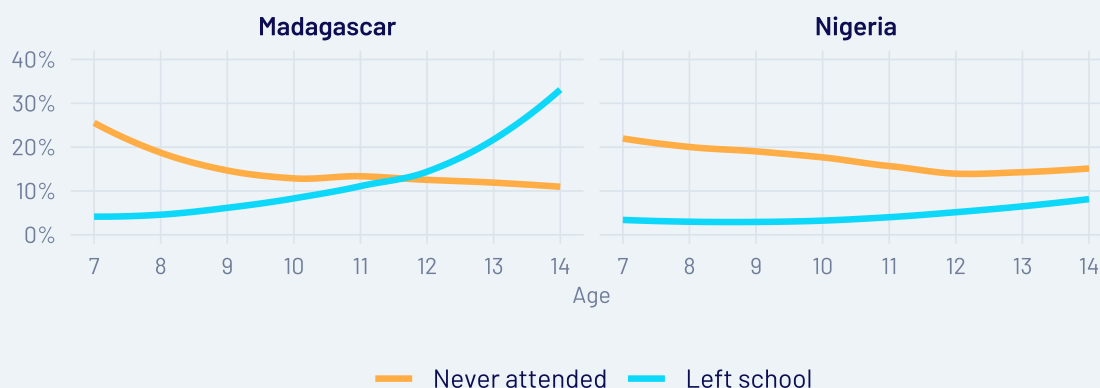
Figure 4.10 illustrates these dynamics using Madagascar and Nigeria, two countries with relatively high levels of non-enrolment throughout childhood.

FIGURE 4.10

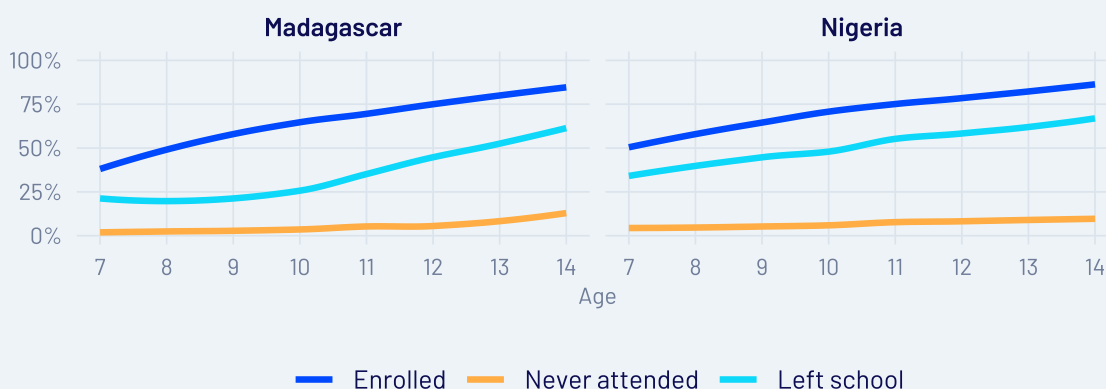
Out-of-school children have very different educational histories and learning outcomes

School participation and mathematics performance by age

Composition of the out-of-school population



Mathematics score by school participation



Notes: Smoothed locally weighted estimates for children aged 7 to 14 in Madagascar and Nigeria. The upper panel shows the proportions of children who have never enrolled and who previously attended school but are currently out of school. The lower panel shows mathematics performance by age separately for children who are currently enrolled, have never enrolled and have left school.
Source: MICS6 Madagascar and Nigeria.

The upper panel shows how the composition of the out-of-school population changes with age. In both countries, the proportion of children who have never attended school declines as children grow older and gradually enter the education system. At the same time, the proportion who have left school increases.

The trajectories differ substantially between the two countries. In Madagascar, dropout increases rapidly from around age eleven, so that by age fourteen most out-of-school children previously attended school. In Nigeria, the opposite pattern is observed. Even by age fourteen, the majority of out-of-school children have never enrolled.

The lower panel compares mathematics performance among children who are enrolled, have never attended



school and have left school.

Unsurprisingly, children who have never attended school record the lowest levels of achievement, with very little improvement across age. Children who have left school perform considerably better than those who never enrolled but remain well behind children who are still attending school.

In Madagascar, the gap between enrolled and out-of-school children narrows during adolescence. This may reflect changes in the composition of the out-of-school population as more children leave school, making the group less selected. Alternatively, older children who leave school may simply have accumulated more years of schooling before dropping out and are therefore better able to complete an assessment pitched at the Grade 2/3 level. Household assessment microdata make it possible to investigate these competing explanations directly.

More broadly, these analyses demonstrate that educational disadvantage is multidimensional. Household wealth influences school participation, school participation influences learning, and both relationships evolve over the course of childhood. Understanding these dynamics requires access to individual-level data rather than summary statistics alone.

WHY MICRODATA MATTER

Educational inequalities cannot be understood solely by comparing average learning outcomes across broad population groups. Publicly available household assessment microdata allow researchers to link learning directly to children's socioeconomic circumstances, educational histories and patterns of school participation. This transforms learning assessments from monitoring tools into resources for understanding how educational disadvantage develops and where interventions may have the greatest impact.

Question 4. When are comparisons between countries valid?

The previous sections demonstrated how publicly available household assessment microdata can improve our understanding of who is represented in assessments, what headline indicators conceal and how learning varies across different groups of children. Comparative analyses, however, require an additional question to be addressed before differences between countries can be interpreted: are the comparisons themselves valid?

International learning assessments are designed to support comparisons across education systems, yet children who are classified as belonging to the same grade often differ in important ways. They may have entered school at different ages, received different amounts of instruction, been assessed at different points in the school year or experienced different language environments. These differences influence measured achievement independently of the quality of the education system itself.

Published assessment reports typically acknowledge some of these issues but rarely investigate their implications empirically. Publicly available microdata make this possible. By combining information on children's age, grade, interview date and schooling history, researchers can identify important sources of non-comparability and develop approaches that improve the interpretation of cross-country comparisons.

This section considers four examples.

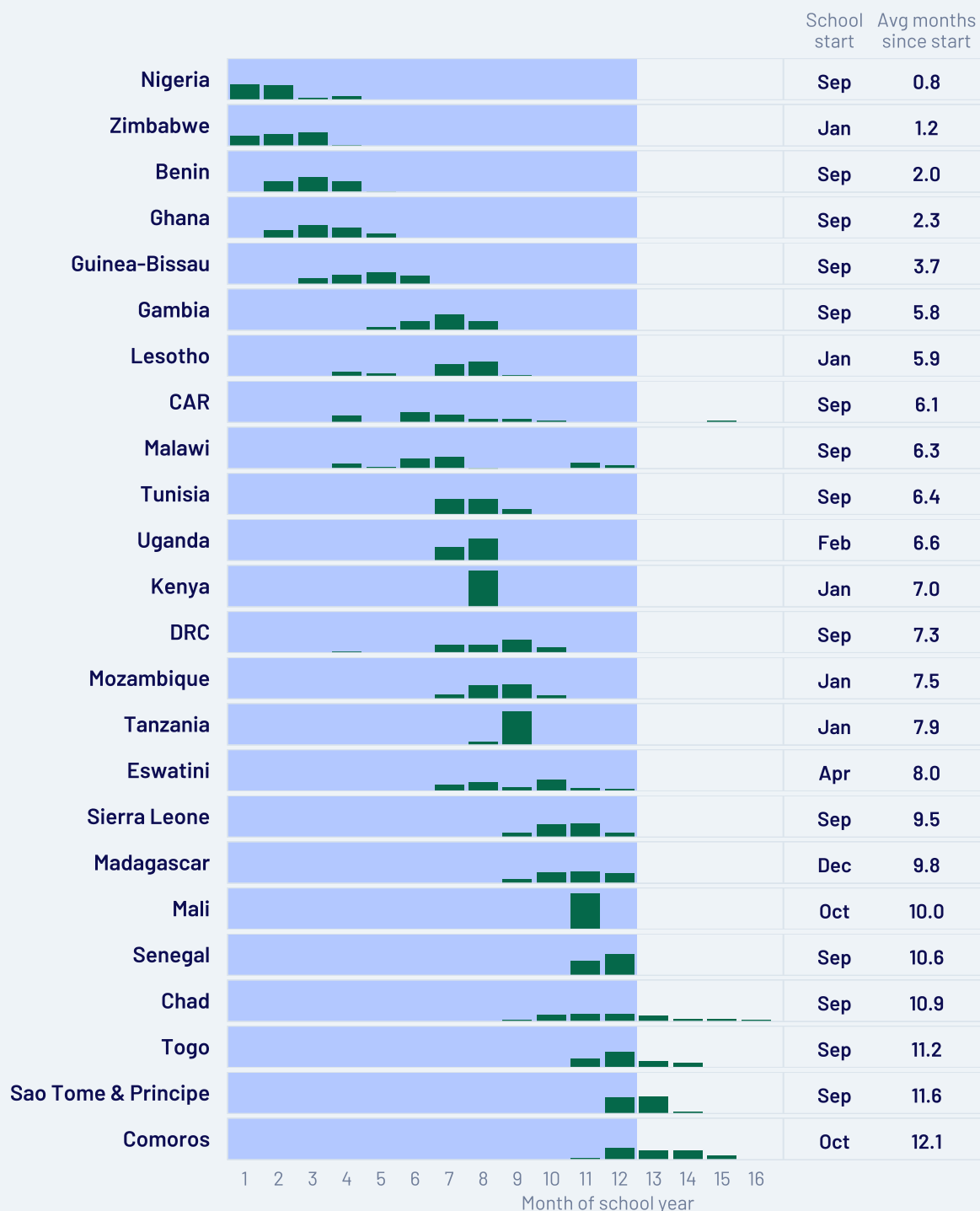
- Are children assessed at comparable points in the school year?
- How does assessment timing influence cross-country comparisons?
- What can learning trajectories tell us about educational progress?
- How do age restrictions affect comparability?

Together these examples demonstrate that rigorous comparative analysis requires more than administering common assessment instruments. It also requires careful attention to the educational experiences represented by the data.

**FIGURE 4.11**

Children in the same grade arrive at the assessment with very different amounts of schooling

Shaded band = one school year



Notes: Distribution of assessments across months since the start of the school year. Excludes months with less than 10 assessments.

Source: MICS6 and ICAN-ICAR 2025 African countries.



Assessment timing and exposure to instruction

One of the least discussed sources of variation in international learning assessments is the timing of data collection. Many comparative analyses implicitly assume that learners in the same grade have received broadly similar amounts of schooling. In practice, this assumption is rarely satisfied. School calendars differ substantially across countries, and assessments may be administered only a few weeks after the school year begins in one country but shortly before the end of the school year in another. Particularly during the early grades, where learning accumulates rapidly within a single academic year, these differences are unlikely to be trivial. Ignoring assessment timing risks attributing differences arising from instructional exposure to differences in educational quality.

Figure 4.11 shows the timing of MICS and ICAN-ICAR fieldwork relative to the beginning of the school year.

Most participating countries begin the school year in September/October. Important exceptions include Zimbabwe, Lesotho, Kenya, Tanzania and Mozambique (January), Uganda (February), Eswatini (April), and Madagascar (December).

Fieldwork also occurs at very different stages of the academic year. Across the MICS surveys, children were assessed on average between approximately one month and eleven months after the beginning of the school year. Zimbabwe represents one extreme, with assessments conducted very early in the school year, whereas children in São Tomé and Príncipe had received almost a full additional year of instruction before being assessed.

The ICAN-ICAR surveys exhibit a different pattern. Fieldwork took place during August in all six participating countries. This corresponds to approximately eight months after the beginning of the school year in Kenya, Tanzania, Uganda and Mozambique, but to the completion of the school year in Mali and Senegal. Correspondence with the PAL Network confirmed that respondents in Mali and Senegal were instructed to report information relating to the school year that had just been completed when answering questions on grade and enrolment.

For the MICS surveys, the timing of assessment relative to the school year was inferred using the constructed variable recording each child's age at the start of the school year. This information was then used to estimate the weighted average number of months of instructional exposure at the time of assessment (final column of Figure 4.11).

These results demonstrate that children assigned to the same administrative grade have often experienced substantially different amounts of schooling when they complete the assessment. Such differences should be considered when interpreting cross-country comparisons.

Adjusting for instructional exposure

Rather than treating all learners in a particular grade as educationally equivalent, it is possible to estimate where children would be expected to perform if they had all been assessed at a common point in the school year.

Figure 4.12 illustrates the effect of adjusting MICS numeracy scores to align all countries to the end of Grade 1. The left-hand point presents observed Grade 1 performance using the original assessment dates. The horizontal placing of the points indicates the point in the school year at which the assessments were completed. The right-hand point adjusts all countries to represent expected performance at the end of Grade 1/ beginning of Grade 2.

Mathematics scores were modelled as a quadratic function of grade level together with the proportion of the school year completed at the time of assessment (e.g. 2.0 represents the beginning of Grade 2, 2.5 the midpoint of Grade 2, and 3.0 the beginning of Grade 3). The model included interaction terms that allowed

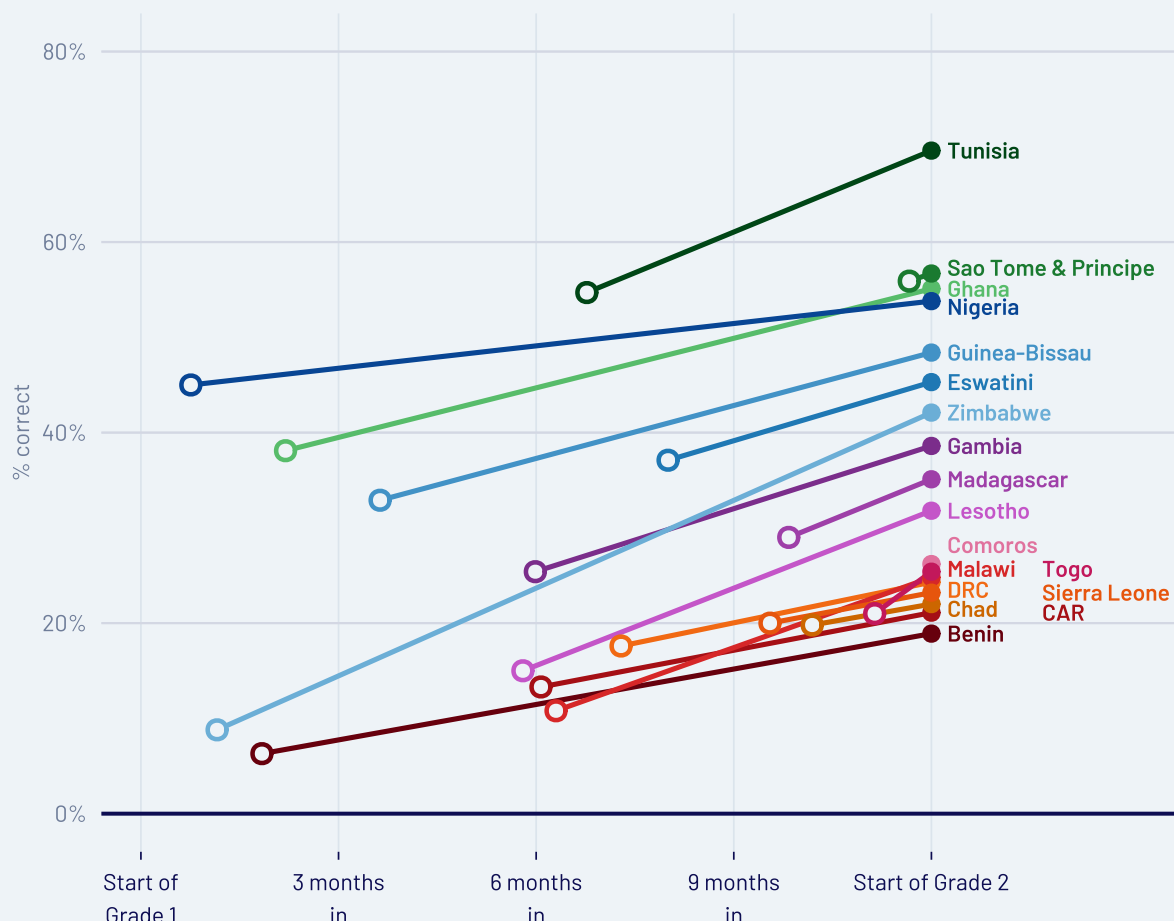


both the level and trajectory of achievement to vary across countries. Model-based predictions therefore represent expected achievement at common points in schooling, independent of the actual timing of data collection in each country³.

FIGURE 4.12

Accounting for instructional exposure can materially change cross-country comparisons

Adjustment of Grade 1 mathematics scores for instructional exposure



Notes: Observed Grade 1 mathematics scores are positioned according to the average point in the school year at which children were assessed. Adjusted estimates represent expected performance at the end of Grade 1/beginning of Grade 2. Mathematics scores are modelled as a quadratic function of grade and the proportion of the school year completed, with country interactions allowing levels and trajectories to vary across countries. The adjustment is intended to illustrate the implications of assessment timing rather than to produce definitive revised country rankings.

Source: AFLEARN analysis of MICS6 African countries.

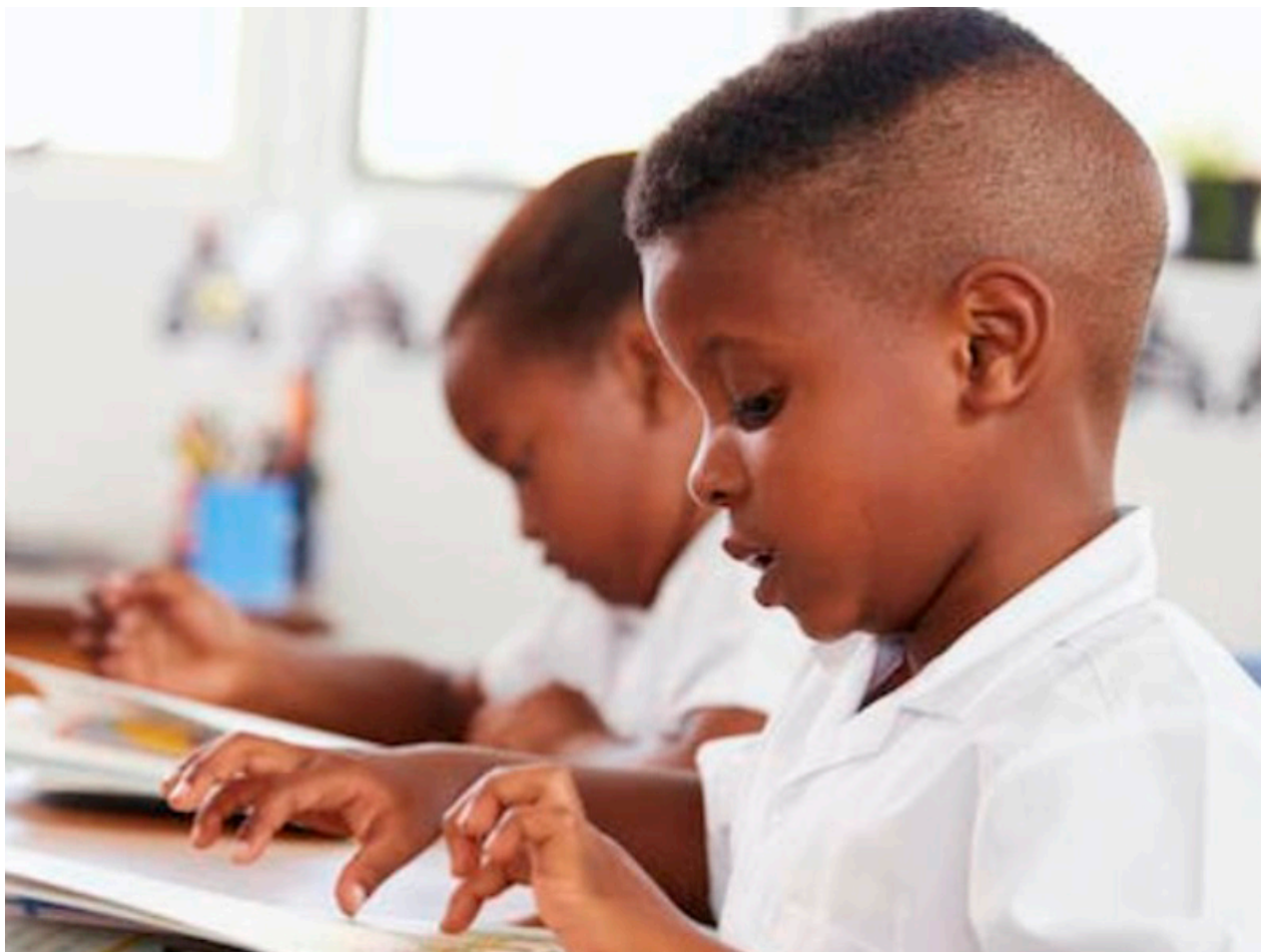
This exercise demonstrates how accounting for instructional exposure can influence cross-country comparisons. The extent of the adjustment depends on both assessment timing and the rate of progression between grades in each country. Several countries change position after adjustment. Zimbabwe, for example, moves from second lowest to seventh once differences in instructional exposure are taken into account.

³ We use a regression model rather than simple interpolation because it facilitates formal statistical tests of differences between countries in both achievement levels and rates of progress while accounting for assessment timing. For descriptive purposes alone, similar adjusted estimates could be obtained by linearly interpolating scores between adjacent grades according to the proportion of the school year completed at the time of assessment.



The purpose of the adjustment is not to produce definitive revised rankings. Rather, it illustrates an important methodological principle. Comparisons based solely on administrative grade implicitly assume equivalent educational experiences where none may exist. Aligning countries according to instructional exposure provides a more meaningful basis for interpreting differences in learning.

Importantly, assessment timing influences not only comparisons at a particular grade but also conclusions about progression across grades. It might be assumed that comparisons of learning growth between grades are unaffected because adjacent grades are always separated by one academic year. However, this assumption overlooks the fact that assessments targeted at Grade 2 and 3 competencies increasingly exhibit ceiling effects in higher grades. As more children approach the upper limit of the assessment, apparent learning gains flatten, even when learning continues. Differences in assessment timing therefore interact with ceiling effects, further complicating comparisons of learning trajectories across countries.



Learning trajectories

The adjustment presented in Figure 4.12 focused on comparisons at a single grade. Assessment timing also influences how we interpret progression across grades. Household assessments administer common instruments across multiple grades and ages, making it possible to estimate learning trajectories. Although these data are cross-sectional rather than longitudinal, they nevertheless provide valuable evidence on how learning accumulates over the primary school years.

These trajectories should be interpreted cautiously. Differences between grades reflect not only learning but also repetition, dropout and selective progression through the education system. They therefore describe patterns of cumulative learning rather than causal estimates of learning growth.

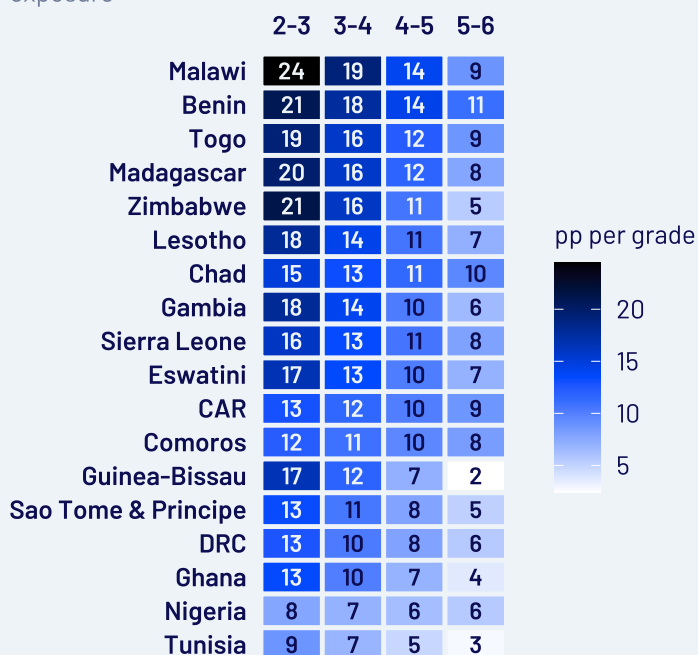


Figure 4.12 also highlighted the considerable variation in mathematics achievement already evident by the end of Grade 1, suggesting that important differences in learning emerge very early in children's school careers. We now turn to consider how the pace of learning differs beyond that point.

FIGURE 4.13

Countries differ markedly in the pace at which mathematics learning accumulates

Learning pace by grade after adjustment for instructional exposure



Notes: Weighted model-based estimates for children enrolled in Grades 1 to 5, adjusted for differences in assessment timing. The assessment is pitched at approximately Grade 2 level; declining estimated learning rates in higher grades therefore partly reflect ceiling effects.

Source: AFLEARN analysis of MICS6 African countries.

Figure 4.13 shows the estimated pace of learning between grades after adjusting for differences in assessment timing. Two findings stand out. First, countries differ markedly in the rate at which learning accumulates. Nigeria, for example, shows relatively modest improvement between Grades 2 and 4, whereas Malawi and Benin exhibit much steeper trajectories. Second, the pace of learning declines progressively across grades in every country. Although some slowing of learning is expected as children progress through school, this assessment is pitched at approximately the Grade 2 level, and part of the observed slowdown reflects ceiling effects as increasing numbers of children approach the upper limit of the assessment.

Countries differ not only in the level of learning they achieve but also in the pace at which learning accumulates. Although there is a degree of persistence in relative performance, countries that begin relatively strongly do not always maintain their advantage and some countries that perform poorly in the early grades do not remain among the weaker performers later in primary school. For example Benin moves from 18th, the lowest rank, at the end of Grade 1 to 8th by the end of Grade 5 while Nigeria moves down from 4th to 13th place. Such differences would remain largely invisible in analyses based on a single grade or a single assessment point.



Comparing mathematics and reading trajectories

The ICAN-ICAR data provide an additional opportunity to compare progression in mathematics and reading within the same education systems.

Figure 4.14 compares the proportion of children reaching the programme-specific proficiency benchmark across grades in the East African countries of Kenya, Tanzania and Uganda.

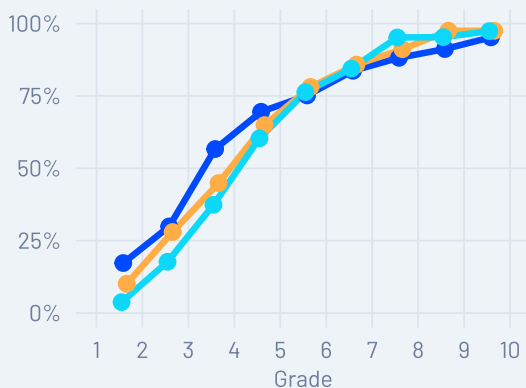
Several findings stand out.

FIGURE 4.14

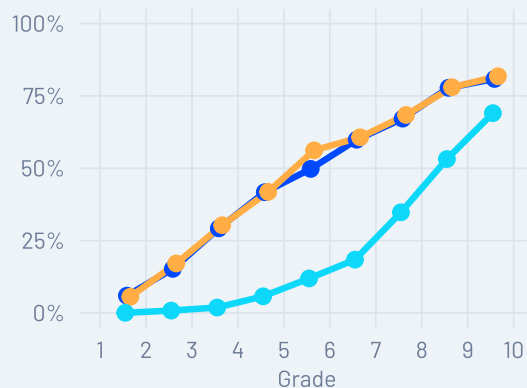
Reading proficiency progresses much more slowly than mathematics, particularly in Uganda

Programme-specific proficiency benchmarks by grade in mathematics and reading

Mathematics



Reading



— Kenya — Tanzania — Uganda

Notes: Weighted estimates for children enrolled in Grades 1 to 9 in Kenya, Tanzania and Uganda. Points are positioned horizontally according to grade and the number of months since the beginning of the school year. Mathematics and reading results use the ICAN-ICAR programme-specific proficiency benchmarks. Source: ICAN-ICAR 2025: Kenya, Tanzania and Uganda.

Across all three countries and at every grade, a larger proportion of children reach the mathematics benchmark than the reading benchmark.

The proportion reaching the mathematics benchmark continues to increase well beyond the Grade 2/3 level at which the benchmark is anchored. By Grade 7, approximately 90% of children in all three countries have acquired Grade 2/3 mathematics proficiency.

Reading follows a markedly different trajectory. Kenya and Tanzania display broadly similar patterns, with around 70% of children reaching the programme-specific reading benchmark by Grade 7. Uganda, however, follows a much flatter progression, with only around 38% reaching the benchmark by the same grade.

The shape of the Ugandan trajectory is particularly informative. Progress between Grades 1 and 3 is relatively limited, followed by more rapid improvement during the upper primary grades. This mirrors the component-level analyses presented in Question 2, where Ugandan children performed especially poorly on listening comprehension and oral reading comprehension despite more encouraging results for basic decoding. Together, these findings suggest that language may play an important role in explaining Uganda's slower progression in reading. This hypothesis is explored further in Question 5.



The value of trajectory analyses therefore lies not simply in comparing levels of achievement but in identifying where learning begins to accelerate, plateau or diverge across countries, thereby generating new questions for investigation.

Age and grade

A final aspect of comparability concerns the relationship between age and grade.

Unlike school-based assessments, household surveys allow learning to be analysed by both age and grade. This provides considerable analytical flexibility but also introduces new comparability challenges because eligibility rules differ across assessment programmes.

The MICS Foundational Learning Skills module is administered only to children aged 7–14, whereas ICAN-ICAR assesses children aged 5–16.

Across the African MICS countries, an average of 15% of Grade 2 learners are excluded because they are younger than seven years old. The proportion varies considerably—from only 3% in Eswatini to 42% in Benin. By Grade 3, the average exclusion falls to approximately 5%.

Given that the assessment is designed to measure competencies expected by the end of the early primary grades, these exclusions are not trivial. Expanding eligibility to include six-year-olds would reduce the proportion of excluded Grade 2 learners to approximately 3%, substantially improving representativeness for analyses focused on the early grades. This would be particularly beneficial where surveys are conducted early in the school year, when younger children are more likely to be enrolled in Grade 2.

This example reinforces a broader methodological lesson. Comparability depends not only on using common assessment instruments but also on ensuring that the populations being compared have similar educational experiences and demographic characteristics.

WHY MICRODATA MATTER

Published assessment reports typically compare countries using administrative grades while paying relatively little attention to differences in instructional exposure, age eligibility or learning trajectories. Publicly available household assessment microdata allow these assumptions to be examined explicitly. Rather than accepting international comparisons at face value, researchers can investigate whether children have experienced comparable amounts of schooling before they are assessed, strengthening both the validity of cross-country comparisons and the interpretation of assessment results.



Question 5. When are language comparisons meaningful?

The previous sections have shown that meaningful comparisons require careful attention to who is represented in the data, how learning is summarised and whether children have experienced comparable amounts of schooling before being assessed. Language introduces an additional—and arguably the most complex—source of variation.

Language is fundamental to both learning and its measurement. Children's opportunities to demonstrate what they know depend not only on their underlying reading or numeracy skills but also on the relationship between the language they speak at home, the language used for instruction and the language in which they are assessed. These relationships differ substantially across African education systems, yet they are rarely examined systematically in comparative assessment reports.

Publicly available household assessment microdata provide a unique opportunity to investigate these relationships directly. Rather than treating language simply as a characteristic of the assessment, the data allow it to be understood as part of children's broader learning environment. By linking information on home language, language of instruction and assessment language at the individual level, researchers can explore how language alignment influences both the interpretation and the comparability of assessment results.

This section considers three related questions.

- What languages are children taught in during the early grades?
- How well do assessment languages align with children's language environments?
- What do these patterns imply for interpreting learning outcomes?

Together, these examples illustrate how publicly available microdata can improve our understanding of one of the most challenging issues in foundational learning assessment.

Language of instruction during the early grades

Before considering assessment language, it is first necessary to understand the language environments in which children are expected to learn. Fewer than one in five African children are taught in their home language, compared with over 70 percent in regions such as South Asia and Latin America. The roots of this misalignment are deeply historical. Colonial languages—English, French, and Portuguese—were elevated to official status, while African languages were excluded from formal systems. Mastery of these official languages often depended on access to formal education, further reinforcing structural inequalities. Although many countries have adopted bilingual or multilingual education policies in more recent decades, implementation has frequently been weak or inconsistent. In some cases, such as Rwanda and Zambia, policy shifts have even reversed prior commitments to mother-tongue instruction.

As a result, the majority of African children continue to be taught and assessed in languages they neither speak at home nor fully understand in the classroom. Understanding the real language of instruction in early grades is difficult, as even where policies exist, they may not be implemented in practice.

Household survey data from MICS help to address this gap by asking children aged 7–14: “What language do your teachers use most of the time when teaching you in class?” This provides a rare glimpse into the de facto language of instruction at scale. ICAN-ICAR asks the adult respondent whether the assessment language is the official medium of instruction in each child's school. While this framing focuses on the official language of instruction, it provides further information on languages used in the classroom in the early grades. The MICS and ICAN-ICAR data is augmented with data from the AMPL assessment in grade 3 in The Gambia and grade 4 in Zambia. AMPL is a school-based survey and principals were asked to provide the primary language of instruction in their school.

Figure 4.15 shows the percentage of children who are taught in colonial languages during the first three years of primary school.

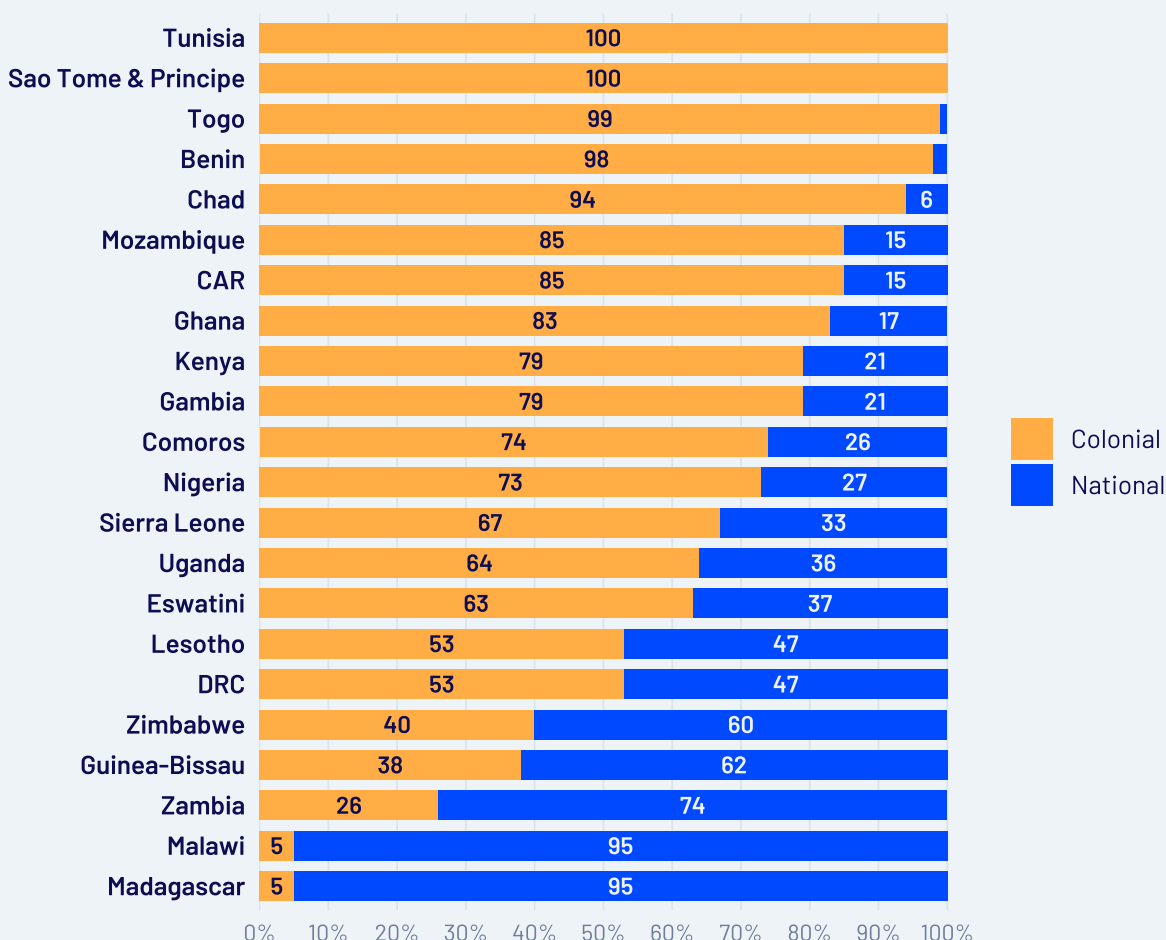
The figure highlights the continued dominance of colonial languages. There are notable exceptions. In Malawi, Madagascar and Zambia, nearly all children are taught in national languages. Francophone countries

generally rely heavily on French from the beginning of primary school, although the Democratic Republic of Congo stands out with 43% of children taught in national languages. Among Lusophone countries, instruction is almost universally in Portuguese in São Tomé and Príncipe and Mozambique, whereas in Guinea-Bissau, about half of learners are taught in Crioulo. There is more diversity in language of instruction across Anglophone countries, although the majority of children in most of these countries are taught in English in the early grades.

FIGURE 4.15

Colonial languages remain the dominant language of instruction in the early grades in most countries

Colonial or National language of instruction in the early grades



Notes: Weighted estimates for children enrolled in Grades 1 to 3. The MICS6 2018 Tunisia survey did not include the question about language of instruction. Therefore the 2023 Tunisia survey is used instead. For ICAN-ICAR, Kenya, Uganda and Mozambique are included; Mali and Senegal are excluded because multiple assessment languages were used, and Tanzania because its sole assessment language was not a colonial language, preventing the relevant question from being used to infer language of instruction. The Grade 4 AMPL sample for Zambia is also included. The Gambia AMPL sample is omitted because both AMPL and MICS estimate that 86% of Grade 3 children are taught in colonial languages.

Source: MICS6, ICAN-ICAR 2025 and AMPL 2023.

These differences mean that children may have experienced very different linguistic environments. Some have been taught almost exclusively in the language they speak at home, while others are already learning through a second, or even third, language.



Language of assessment

Assessment programmes adopt different approaches to multilingual education systems, with the choice of assessment language typically determined by countries in line with national policy and operational considerations. Table A1 in the Appendix summarises the languages of assessment used by the major assessment programmes in each country. Although most early-grade assessments are administered in colonial languages, several programmes include widely spoken African languages or are conducted exclusively in the national language of instruction. In MICS6, multiple assessment languages are available in Eswatini, Lesotho, Madagascar, Malawi, Nigeria and Zimbabwe. ICAN-ICAR uses multiple assessment languages in Mali and Senegal and is administered exclusively in Swahili in Tanzania. In the 2019 round of PASEC, Grade 2 learners in Burundi and Madagascar were assessed in Kirundi and Malagasy respectively. The use of African languages expanded further in the 2024 round, with multiple assessment languages introduced in Mali, Niger and Senegal. Similarly, PIRLS in South Africa is administered in all 11 official languages.

Aligning children's language environments

Understanding language in education requires moving beyond simple descriptions of the language of instruction or the language of assessment. What matters is how the different components of a child's language environment fit together.

Both MICS and AMPL ask children about the language they speak most frequently at home. MICS asks, "Which language do you speak most of the time at home?", while AMPL asks, "What is the main language that you speak at home?" ICAN-ICAR adopts a slightly different approach, asking the adult respondent whether the assessment language is spoken at home and, if so, how frequently. In the analyses presented here, the assessment language is treated as a home language only when respondents report that it is spoken always, rather than sometimes or never.

Combining these responses with information on the language of instruction and the language used for assessment makes it possible to describe children's language environments consistently across assessment programmes.

Figure 4.16 shows the alignment between home language, language of instruction and assessment language.

Several broad patterns emerge.

At one end of the spectrum are countries where children are assessed in the same language they speak at home and the same language used for classroom instruction. This represents the strongest alignment and provides the clearest basis for interpreting assessment results as measures of children's underlying competencies.

A second group of countries shows strong alignment between the language of instruction and the language of assessment, even though many children speak different languages at home. This pattern characterises several Francophone countries, where French functions as the dominant language of instruction throughout the early grades. Although the assessment language may not coincide with children's home language, it nevertheless reflects the language through which formal learning takes place.

The greatest concerns arise where children are assessed in a language that is neither their home language nor their language of instruction. In these circumstances, poor assessment performance may reflect limited familiarity with the assessment language rather than underlying reading or numeracy skills, making the interpretation of results considerably more difficult. In Zambia the AMPL assessment is conducted exclusively in English, despite the fact that English is only the language of instruction in 24% of the schools. In the Democratic Republic of the Congo and Guinea-Bissau, both MICS and PASEC assess children in French and Portuguese, despite substantial use of other languages during instruction in the early grades.

For the MICS reading assessment, children were generally not assessed if the available assessment languages matched neither their reported home language nor their language of instruction. Consequently, language mismatch contributes directly to the missing data observed in the MICS reading assessment.

These examples demonstrate that language alignment varies considerably across both countries and assessment programmes. Describing the language of assessment alone is therefore insufficient. Understanding children's broader language environments is essential for interpreting learning outcomes.

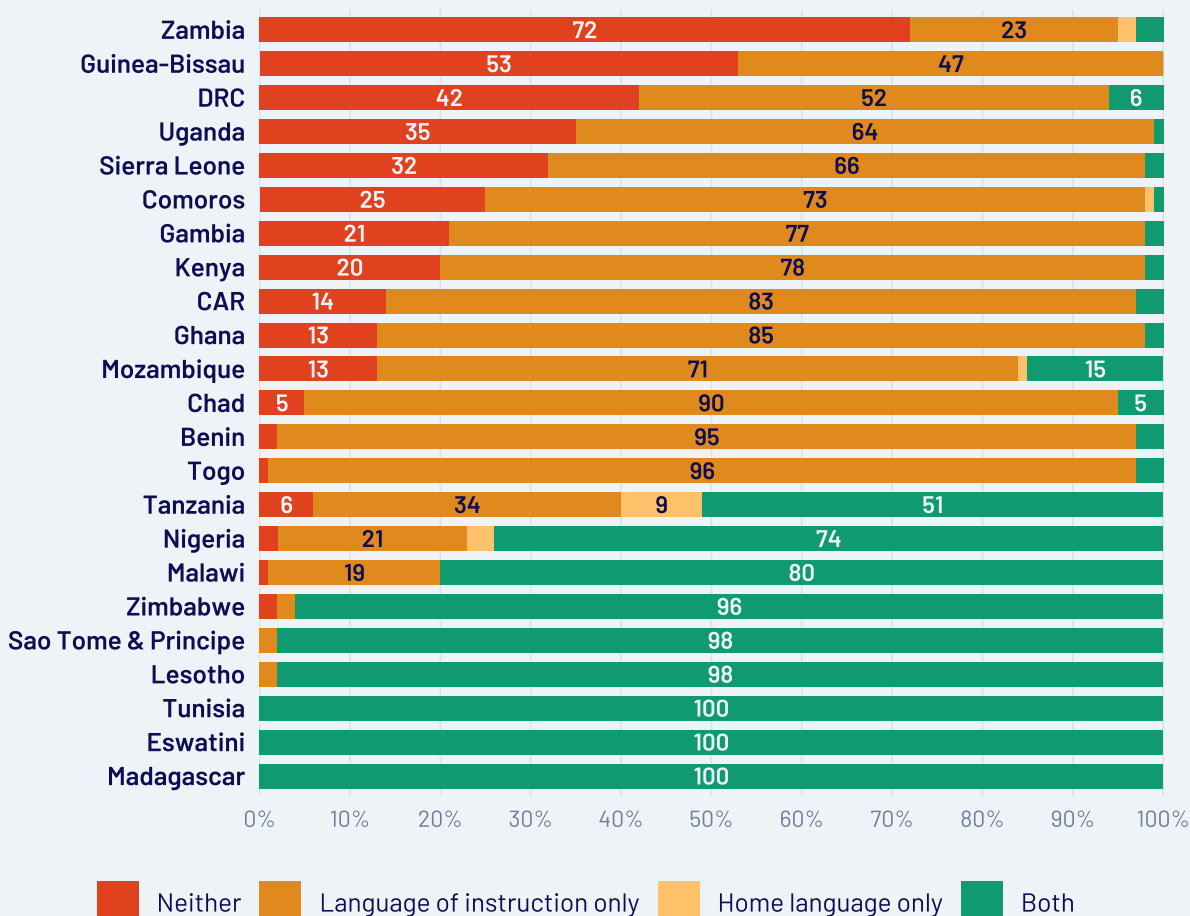
Language and the interpretation of learning outcomes

The importance of these language pathways is illustrated by several of the analyses presented earlier in this chapter.

FIGURE 4.16

Assessment languages vary greatly in how well they align with children's home and instructional languages

Alignment between home language, language of instruction and assessment language



Notes: Weighted estimates for children enrolled in Grades 2 and 3. Categories indicate whether the assessment language is also the child's home language, language of instruction, both, or neither. For MICS, Tunisia 2023 is used because Tunisia 2018 did not include the language-of-instruction question. For ICAN-ICAR, Kenya, Uganda, Mozambique and Tanzania are included; Mali and Senegal are excluded because the household-language question does not permit equivalent classification when multiple assessment languages are used. The Zambian Grade 4 and Gambian Grade 3 AMPL samples are included. For ICAN-ICAR, the assessment language is classified as a home language only when respondents report that it is always spoken at home.

Source: MICS6, ICAN-ICAR 2025 and AMPL 2023.



Question 1 showed that language mismatch contributes substantially to lower realised response rates for the MICS reading assessment because many children are ineligible to attempt the reading passage when suitable assessment languages are unavailable. Consequently, the realised reading sample differs systematically from both the intended survey population and the realised mathematics sample.

Question 2 demonstrated how analysing assessment components helps identify where reading development begins to break down. In Uganda, for example, children performed relatively well on letter-sound recognition but very poorly on listening comprehension and oral reading comprehension. These patterns suggest that limitations in oral language development may constrain reading comprehension even where basic decoding skills have begun to develop.

Question 4 showed that Uganda also exhibits a much flatter trajectory towards the programme-specific reading proficiency benchmark than Kenya and Tanzania, despite relatively similar progression in mathematics. Taken together, these findings suggest that language may represent an important dimension of cross-country differences in reading achievement, although they do not in themselves establish a causal explanation.

This illustrates a broader methodological principle. Language variables should not be treated simply as contextual background information. They are integral components of comparative analysis that help researchers distinguish between differences arising from children's learning and differences arising from the linguistic contexts in which learning is assessed.

Looking forward

The analyses presented here also point towards opportunities for improving future assessments.

The feasibility of assessing children in African languages has already been demonstrated. Early Grade Reading Assessments (EGRAs), for example, have been adapted into more than ninety African languages. While many of these assessments were implemented as part of donor-funded projects, nationally representative examples from countries such as South Africa and Kenya demonstrate that large-scale multilingual assessment is technically feasible.

The principal barriers are therefore no longer technical but institutional. They include the costs of developing and maintaining multiple language versions, limited technical capacity, complex language policies and the understandable desire to maximise comparability across countries. As foundational learning assessments continue to evolve, however, improving the alignment between children's language environments and assessment languages offers one of the clearest opportunities for strengthening both validity and equity.

More generally, language exemplifies a broader theme running throughout this report. Producing more comparable assessments does not necessarily require making assessments more uniform. Rather, it requires understanding the contexts in which learning occurs and ensuring that differences in assessment outcomes are interpreted appropriately.

WHY MICRODATA MATTER

Published assessment reports typically describe the language of assessment but provide limited opportunities to investigate how children's language environments influence learning outcomes. Publicly available household assessment microdata make it possible to link home language, language of instruction and assessment language at the individual level, allowing researchers to investigate one of the most important sources of variation in foundational learning assessment. In doing so, they strengthen not only the interpretation of assessment results but also the design of future assessments.

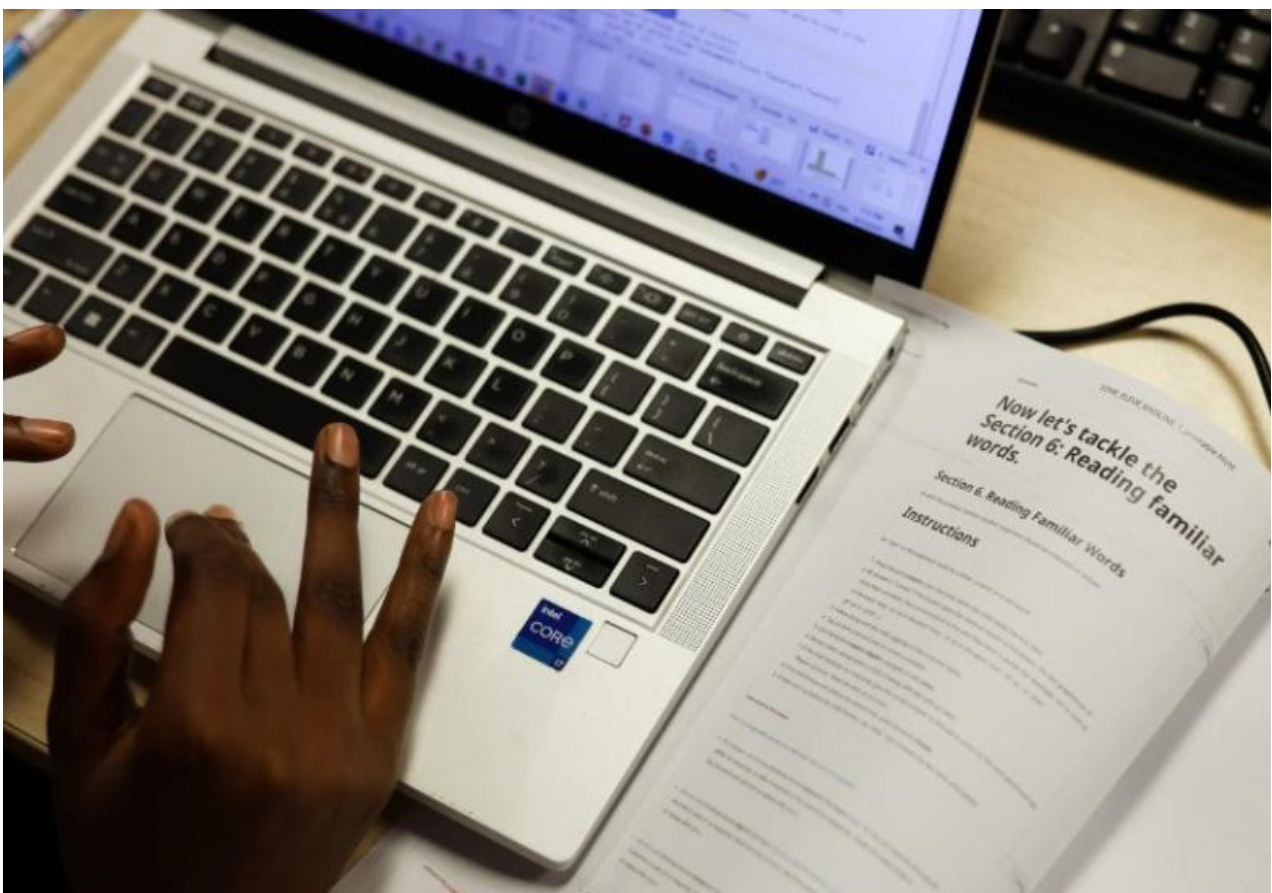
4.3 Chapter conclusion

The analyses presented in this chapter demonstrate that the value of publicly available household assessment microdata extends far beyond producing additional statistics or replicating published reports. By making individual-level data accessible, they allow researchers to investigate fundamental questions about how learning assessments should be interpreted, compared and improved.

Across the five questions considered in this chapter, a common theme has emerged. Publicly available microdata enable researchers to move beyond describing learning outcomes towards evaluating the assumptions that underpin those outcomes. They reveal who is represented in assessment samples, what information is hidden by headline indicators, how educational inequalities emerge, when comparisons between countries are valid and how children's language environments influence the interpretation of learning.

The examples presented here are necessarily illustrative rather than exhaustive. Every publicly available assessment dataset contains opportunities for further methodological investigation. The greatest return on investment in foundational learning assessment therefore comes not only from collecting high-quality data but also from ensuring that those data remain openly accessible for rigorous secondary analysis.

As Africa's foundational learning assessment ecosystem continues to mature, the combination of high-quality assessment programmes, open microdata and rigorous secondary analysis offers an increasingly powerful foundation for improving both educational evidence and educational policy. Open data do more than promote transparency and reproducibility. They improve the interpretation, comparability and quality of foundational learning assessments themselves.





Chapter 5: From Assessment Data to Better Evidence

Strengthening Africa's Foundational Learning Evidence Ecosystem

Key message

Africa has made substantial investments in measuring foundational learning over the past two decades. National, regional and international assessment programmes now provide an increasingly rich picture of children's reading and numeracy across diverse educational contexts. Yet collecting assessment data is only the beginning of the evidence journey. The long-term value of these investments depends on whether assessment data are documented, preserved, made accessible and actively reused. Strengthening every stage of this evidence ecosystem—from assessment design and data stewardship to rigorous secondary analysis—will increase the return on existing investments while improving future assessments, educational research and policymaking.

5.1 From assessment data to better evidence

This report has argued that strengthening Africa's foundational learning evidence ecosystem requires attention to much more than the collection of new assessment data.

Chapter 1 introduced the concept of the foundational learning evidence ecosystem and argued that sustainable evidence depends on three interconnected layers: high-quality assessment programmes, a robust data value chain, and a broader ecosystem of institutions, partnerships and capacity that supports the long-term production and use of evidence.

Chapter 2 demonstrated that Africa has experienced a remarkable expansion in foundational learning assessment over the past two decades. National, regional and international programmes now cover much of the continent and provide increasingly diverse sources of information on children's learning. At the same time, the chapter highlighted continuing challenges relating to comparability, language, documentation and uneven geographical coverage, illustrating that assessment systems continue to evolve.

Chapter 3 shifted attention from assessment programmes themselves to the data they generate. While access to microdata has improved substantially in recent years—with important advances such as the public release of PASEC microdata, the availability of ICAN-ICAR datasets through DataFirst and the recovery of more than two decades of EGRA datasets—significant gaps remain. Many nationally commissioned assessments are still inaccessible, documentation varies widely between programmes, and important datasets remain at risk of being permanently lost. The chapter argued that applying the FAIR principles provides a practical framework for transforming assessment data into long-term public assets.

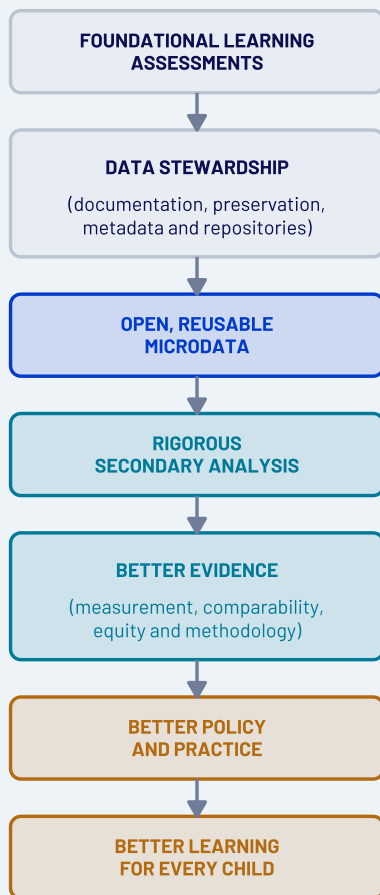
Chapter 4 demonstrated why this transformation matters. Through five methodological questions, it showed that publicly available microdata allow researchers to move beyond headline indicators to investigate representation, examine learning distributions, understand educational inequalities, strengthen cross-country comparability and explore how children's language environments influence assessment results. These analyses illustrate that open microdata do more than increase transparency—they improve the interpretation of assessment evidence itself.

Taken together, the four chapters point to a common conclusion. Africa's next challenge is not simply to produce more assessment data. It is to ensure that the substantial investments already made continue to generate evidence long after fieldwork has been completed.

FIGURE 5.1

Better evidence depends on strengthening the entire assessment data ecosystem

From assessment data to better evidence



The greatest return on investment comes not simply from collecting assessment data but from ensuring that they remain accessible, reusable and actively used.

● Investment ● Enduring asset ● New value ● Outcomes

5.2 Strengthening the evidence ecosystem

Strengthening Africa's foundational learning evidence ecosystem requires coordinated action across the entire data value chain. Different institutions have distinct responsibilities, but they share a common objective: maximising the long-term value of assessment data while maintaining high standards of quality, ethics and public accountability.

The recommendations below are organised according to the stages of the evidence ecosystem rather than by institution.



Recommendation 1: Design assessments for long-term reuse

The process of creating reusable assessment data begins before the first child is assessed.

Assessment programmes should integrate documentation, metadata creation, anonymisation and microdata dissemination planning into assessment design rather than treating these as activities to be undertaken after national reports have been published. Governments should establish clear governance arrangements specifying responsibilities for long-term stewardship, preservation and access, ensuring that data sharing is planned from the outset.

Assessment programmes should also routinely document variables shown in Chapter 4 to influence interpretation and comparability, including assessment timing, realised response rates, age eligibility, school sector coverage, language of instruction and language of assessment.

Recommendation 2: Preserve and document Africa's assessment heritage

Africa has already generated an significant volume of foundational learning assessment data. Much of this evidence remains underused, while some historical datasets are at risk of permanent loss.

Governments and development partners should invest in recovering legacy datasets, preserving historical documentation and developing consistent metadata standards across assessment programmes. Documentation—including questionnaires, codebooks, sampling reports, technical manuals and syntax files—should be regarded as an essential assessment output rather than supplementary material.

The recovery of more than 2,500 USAID-funded EGRA files described in Chapter 3 demonstrates both the scale of the challenge and the feasibility of preserving valuable historical evidence before it disappears.

Recommendation 3: Make assessment microdata as open as possible

Publicly available microdata increase the scientific, policy and public value of assessment investments.

Governments and assessment organisations should adopt FAIR-aligned data governance frameworks that make datasets as open as possible while protecting confidentiality and respecting legal and ethical obligations. Trusted repositories should become the default mechanism for preserving and disseminating assessment data, ensuring that datasets remain discoverable, accessible and reusable over time.

Open access should extend beyond data files to include comprehensive documentation, standardised metadata and supporting analytical resources that lower barriers to reuse.

Recommendation 4: Build African capacity for secondary analysis

Making data available is only the first step. The value of open data depends on whether researchers, governments and practitioners have the capacity to use them effectively.

Universities, research institutions and assessment organisations should strengthen postgraduate training, short courses and professional development focused on secondary analysis of foundational learning data. Investments should extend beyond statistical methods to include data stewardship, reproducible research, measurement, psychometrics and the responsible use of emerging AI tools to support data analysis.

Developing open analytical code, practical guides and multilingual training materials can further reduce technical barriers and broaden participation in comparative research.

Recommendation 5: Strengthen methodological research

One of the central findings of Chapter 4 is that publicly available microdata allow researchers not only to analyse learning outcomes but also to improve the assessments themselves.

Future research should continue investigating issues such as instructional exposure, language pathways, realised samples, assessment timing, comparability across programmes and the interpretation of programme-



specific proficiency benchmarks. These analyses strengthen the evidence base while informing future assessment design.

Assessment organisations should therefore view secondary analysis as an integral part of assessment development rather than an activity that begins only after reporting has been completed.

Recommendation 6: Build a collaborative evidence ecosystem

No single organisation can create a strong foundational learning evidence ecosystem on its own.

Governments, assessment organisations, universities, regional bodies and development partners each contribute different forms of expertise, infrastructure and capacity. Continued progress will depend upon strengthening collaboration across these institutions while promoting common documentation standards, interoperable metadata and trusted regional repositories.

Regional collaboration also provides opportunities to harmonise analytical resources, coordinate capacity building and support comparative research that explicitly addresses issues of comparability rather than assuming them.

5.3 The role of AFLEARN

The African Foundational Learning Data Hub (AFLEARN) was established to strengthen the evidence ecosystem rather than to replace existing assessment programmes.

Its role spans the entire data value chain.

AFLEARN supports the preservation and curation of assessment datasets; develops harmonised documentation and analytical resources; improves access through trusted repositories; strengthens African capacity through training and mentoring; undertakes methodological research on assessment quality and comparability; and provides technical advice to governments, assessment organisations and development partners.

In doing so, AFLEARN connects institutions that traditionally operate at different stages of the evidence ecosystem. Rather than generating new assessments itself, it increases the value of existing investments by ensuring that assessment data remain accessible, understandable and actively used.

As artificial intelligence increasingly transforms how researchers interact with data, AFLEARN is also well positioned to support African institutions in adopting these technologies responsibly. Its expertise in data stewardship, documentation, open science and methodological research provides a strong foundation for developing practical guidance on the ethical and effective use of AI in the analysis of foundational learning data.

5.4 Conclusion

Africa has already demonstrated that it can design and implement high-quality foundational learning assessments across diverse educational contexts. The next challenge is to ensure that these investments continue to generate value long after fieldwork has ended. Treating assessment data as enduring public assets—supported by robust stewardship, open access and rigorous secondary analysis—will increase the return on every assessment undertaken. In doing so, Africa can strengthen not only its evidence base but also its capacity to improve learning opportunities for future generations of children.



References

- Advocates for International Development. (2022). SDG 4: Quality education—A legal guide. Advocates for International Development. https://www.right-to-education.org/sites/right-to-education.org/files/resource-attachments/A4ID_SDG%20Legal%20Guide_SDG4_Quality%20education_February2022_EN.pdf
- Evans, D. K., & Mendez Acosta, A. (2023). How to measure student absenteeism in low- and middle-income countries. *Economics of Education Review*, 96, Article 102454. <https://doi.org/10.1016/j.econedurev.2023.102454>
- Gahutu, P. (2024). Rwanda's Language Policy from the Past to the Present Day. In E. M. Lisanza & L. Muaka (Eds.), *The Palgrave Handbook of Language Policies in Africa* (pp. 131- 148). Springer International Publishing. https://doi.org/10.1007/978-3-031-57308-8_759
- International Association for the Evaluation of Educational Achievement. (n.d.). Progress in International Reading Literacy Study (PIRLS). Retrieved July 21, 2026, from <https://www.iea.nl/studies/iea/pirls>
- PAL Network. (n.d.). Datasets. Retrieved July 21, 2026, from <https://palnetwork.org/datasets/>
- PAL Network. (n.d.). Literacy & numeracy for every child (ICAN–ICAR). Retrieved July 21, 2026, from <https://palnetwork.org/ican-icar/>
- People's Action for Learning Network. (2026). International Common Assessment of Numeracy- International Common Assessment of Reading 2025 (Version 1) [Data set]. DataFirst. <https://doi.org/10.25828/AH42-XW77>
- TIMSS & PIRLS International Study Center. (n.d.). Databases. Boston College. Retrieved July 21, 2026, from <https://timssandpirls.bc.edu/databases-landing.html>
- Trudell, B. (2024). Language in education policy: Key issues and their enactment in The Gambia. UNESCO. <https://doi.org/10.54676/GPJH5598>
- UNESCO Institute for Statistics, & Australian Council for Educational Research. (2024). Assessments for Minimum Proficiency Levels a and b (AMPL-ab): International report. UNESCO Institute for Statistics. <https://doi.org/10.54675/ELPT8720>
- UNESCO Institute for Statistics. (n.d.). AMPL-ab project. Retrieved July 21, 2026, from <https://www.uis.unesco.org/en/methods-and-tools/ampl/ampl-ab-project>
- UNESCO. (2024). Spotlight report on basic education completion and foundational learning in Africa: Learning counts. UNESCO, Association for the Development of Education in Africa (ADEA), & African Union. <https://doi.org/10.54676/TBRB1563>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018.



Data sources

- Bolgrien, A., Boyle, E. H., Sobek, M., & King, M. (2026). IPUMS MICS data harmonization code (Version 1.5) [Stata syntax]. IPUMS: Minneapolis, MN. <https://doi.org/10.18128/D082.V1.6>
- Bureau of Statistics and UNICEF. (2018). Lesotho Multiple Indicator Cluster Survey 2018 [Dataset]. Bureau of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- Central Statistical Office and UNICEF. (2021). Eswatini Multiple Indicator Cluster Survey 2021–2022 [Dataset]. Central Statistical Office and UNICEF. Retrieved from <https://mics.unicef.org>
- CONFEMEN/PASEC. (2026). PASEC 2019 International Database [Dataset]. Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN. Retrieved from <https://pasec.confemen.org/en/data-portal>
- Gambia Bureau of Statistics and UNICEF. (2018). The Gambia Multiple Indicator Cluster Survey 2018 [Dataset]. Gambia Bureau of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- Ghana Statistical Services and UNICEF. (2017). Ghana Multiple Indicator Cluster Survey 2017–2018 [Dataset]. Ghana Statistical Services and UNICEF. Retrieved from <https://mics.unicef.org>
- Institute of Statistics, Demographics, and Economics and UNICEF. (2019). Chad Multiple Indicator Cluster Survey 2019 [Dataset]. Institute of Statistics, Demographics, and Economics and UNICEF. Retrieved from <https://mics.unicef.org>
- Institute of Statistics, Economics, and Social Studies and UNICEF. (2018). Central African Republic Multiple Indicator Cluster Survey 2017–2018 [Dataset]. Institute of Statistics, Economics, and Social Studies and UNICEF. Retrieved from <https://mics.unicef.org>
- Ministry of Development and International Cooperation and National Institute of Statistics and UNICEF. (2018). Tunisia Multiple Indicator Cluster Survey 2018 [Dataset]. Ministry of Development and International Cooperation and National Institute of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- Ministry of Development and International Cooperation and National Institute of Statistics and UNICEF. (2023). Tunisia Multiple Indicator Cluster Survey 2023 [Dataset]. Ministry of Development and International Cooperation and National Institute of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- Ministry of Economy and the State Secretariat for Planning and Regional Integration and UNICEF. (2018). Guinea-Bissau Multiple Indicator Cluster Survey 2018–2019 [Dataset]. Ministry of Economy and the State Secretariat for Planning and Regional Integration and UNICEF. Retrieved from <https://mics.unicef.org>
- National Bureau of Statistics and UNICEF. (2021). Nigeria Multiple Indicator Cluster Survey 2021 [Dataset]. National Bureau of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- National Institute of Statistics (INSEED) and UNICEF. (2017). Togo Multiple Indicator Cluster Survey 2017 [Dataset]. National Institute of Statistics (INSEED) and UNICEF. Retrieved from <https://mics.unicef.org>
- National Institute of Statistics and Economic Analysis and UNICEF. (2021). Benin Multiple Indicator Cluster Survey 2021–2022 [Dataset]. National Institute of Statistics and Economic Analysis and UNICEF. Retrieved from <https://mics.unicef.org>
- National Institute of Statistics and UNICEF. (2017). Democratic Republic of the Congo Multiple Indicator Cluster Survey 2017–2018 [Dataset]. National Institute of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- National Institute of Statistics and UNICEF. (2018). Madagascar Multiple Indicator Cluster Survey 2018 [Dataset]. National Institute of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- National Institute of Statistics and UNICEF. (2019). São Tomé and Príncipe Multiple Indicator Cluster Survey 2019 [Dataset]. National Institute of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- National Institute of Statistics and UNICEF. (2022). Comoros Multiple Indicator Cluster Survey 2022 [Dataset]. National Institute of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>



- National Office of Statistics and UNICEF. (2018). Algeria Multiple Indicator Cluster Survey 2018–2019 [Dataset]. National Office of Statistics and UNICEF. Retrieved from <https://mics.unicef.org>
- National Statistical Office and UNICEF. (2019). Malawi Multiple Indicator Cluster Survey 2019–2020 [Dataset]. National Statistical Office and UNICEF. Retrieved from <https://mics.unicef.org>
- National Statistics Agency and UNICEF. (2019). Zimbabwe Multiple Indicator Cluster Survey 2019 [Dataset]. National Statistics Agency and UNICEF. Retrieved from <https://mics.unicef.org>
- People's Action for Learning Network. (2026). International Common Assessment of Numeracy- International Common Assessment of Reading 2025 (Version 1) [Data set]. DataFirst. <https://doi.org/10.25828/AH42-XW77>
- Statistics Sierra Leone and UNICEF (2017). Sierra Leone Multiple Indicator Cluster Survey 2017 [Dataset]. Statistics Sierra Leone and UNICEF. Retrieved from <https://mics.unicef.org>
- UNESCO Institute for Statistics. (2024). Assessment of Minimum Proficiency Levels (AMPL-ab) [Data set]. Retrieved from <https://www.uis.unesco.org/en/methods-and-tools/ampl/ampl-ab-project>

**TABLE A1. Assessment language by country and programme**

Country	Assessment language	Assessments (most recent)
Benin	French	MICS6 & PASEC 2024
Botswana	English	prePIRLS 2011
Burkina Faso	French	PASEC 2024
Burundi	Kirundi	PASEC 2024
Cameroon	French, English	PASEC 2024
Central African Republic	French	MICS6 & PASEC 2024
Chad	French	MICS6 & PASEC 2024
Comoros	French	MICS6
Congo	French	PASEC 2024
Cote d'Ivoire	French	PASEC 2024
Democratic Republic of Congo	French	MICS6 & PASEC 2024
Djibouti	French	PASEC 2024
Eswatini	English, Siswati	MICS6
Gabon	French	PASEC 2024
Gambia	English	MICS6 & AMPL 2023
Ghana	English	MICS6
Guinea	French	PASEC 2024
Guinea Bissau	Portuguese	MICS6 & PASEC 2024
Kenya	English	ICAN-ICAR 2025
Lesotho	English, Sesotho	MICS6
Madagascar	French, Malagasy	MICS6
Malawi	English, Chichewa	MICS6
Mali	French, Bamanankan	ICAN-ICAR 2025 & PASEC 2024
Mozambique	Portuguese	ICAN-ICAR 2025
Niger	French, Haoussa, Zarma	PASEC 2024
Nigeria	English, Hausa, Igbo, Yoruba	MICS6
Nigeria	English	PASEC 2024
Sao Tome	Portuguese	MICS6 & PASEC 2024
Senegal	French, Wolof, Soninke, Sereere, Pulaar, Mandinka, Joola, Diola	ICAN-ICAR 2025 & PASEC 2024
Sierra Leone	English	MICS6
South Africa	English, Afrikaans, isiZulu, isiXhosa, Sepedi, Setswana, Xitsonga, Sesotho, Tshivenda, SiSwati, isiNdebele	PIRLS 2021
Tanzania	Swahili	ICAN-ICAR 2025
Togo	French	MICS6 & PASEC 2024
Uganda	English	ICAN-ICAR 2025
Zambia	English	AMPL 2023
Zimbabwe	English, Ndebele, Shona	MICS6



African Foundational Learning Data Hub

DataFirst

University of Cape Town

www.datafirst.uct.ac.za/aflearn