MICS

Toward achieving inclusive and equitable quality education for all

A manual for statistical data analysis using Multiple Indicator Cluster Surveys (MICS6) with a special focus on achieving the Sustainable Development Goals

unicef | for every child

ACKNOWLEDGEMENTS

This manual for statistical data analysis using Multiple Indicator Cluster Surveys (MICS) was developed by Diogo Amaro, under the supervision of Suguru Mizunoya, in the Data and Analytics section, Division of Data, Analytics, Planning and Monitoring.

We would like to thank the Bill and Melinda Gates Foundation for financing the initial stages of the MICS-EAGLE initiative, which included the production of this manual, through the Strategic Grant Fund provided to UNICEF.

We would like to express our gratitude to our colleagues from UNICEF's MICS team (Ivana Bjelic, Yadigar Coskun, Attila Hancioglu, Glen Heller, Shane Khan, Bo Pedersen, Eva Quintana and Turgay Unalan) and at UNESCO's Institute for Statistics (Friedrich Huebler, Alasdair McWilliam and Silvia Montoya) for their inputs to early drafts of the manual.

We would also like to thank Takako Shimizu, Xinxin Yu and our other colleagues in UNICEF's Programme Division and Regional Offices, without whose support this initiative could not have advanced, and Sakshi Mishra and Haogen Yao who supported the review and production of the codes for the digital annex.

And, we would like to express our appreciation of colleagues who envisaged and developed innovative education modules for MICS6, including Gabrielle Bonnet, Jo Bourne, Matt Brossard, Manuel Cardoso, Hiroyuki Hattori, Daniel Kelly, Blandine Ledoux and Francis Ndem.

Last but not least, the team would like to thank Anna Giovinetto for editing the manual and Cecilia Silva Venturini for its design.

CONTACT

Suguru Mizunoya, Statistics & Monitoring Specialist
smizunoya@unicef.org
Data and Analytics Section
United Nations Children's Fund
3 United Nations Plaza, New York, NY 10017, USA

MICS

Toward achieving inclusive and equitable quality education for all

A manual for statistical data analysis using Multiple Indicator Cluster Surveys (MICS6) with a special focus on achieving the Sustainable Development Goals

CONTENTS

ACRONYMS

ANAR	Adjusted Net Attendance Rate
DHS	Demographic and Health Surveys
ECDI	Early Child Development Index
ECE	Early Childhood Education
ICT	Information And Communication Technology
ISCED	International Standard Classification of Education
ITU	International Telecommunication Union
Lao PDR	Lao People's Democratic Republic
MICS	Multiple Indicator Cluster Survey
MICS-EAGLE	MICS Education Analysis for Global Learning and Equity
SDG	Sustainable Development Goal
UIS	UNESCO Institute for Statistics
UNSD	United Nations Statistics Division
WIDE	World Inequalities Database on Education

INTRODUCTION

This guide was prepared by the Education team of UNICEF's Data and Analytics section. It details the calculation process for various education indicators and is positioned as a helpful tool for developing national governments' statistical capacity, helping these stakeholders understand how to calculate and use key education indicators and link them to policy discussions. The manual is freely available online, a cost-effective measure that ensures it can be used by government staff and researchers around the world to promote better statistical evidence for impactful action.

Part of the broader Multiple Indicator Cluster Surveys-Education Analysis for Global Learning and Equity (MICS-EAGLE) initiative, this guide was developed to help countries better understand the status of their education systems and make informed decisions based on data-driven evidence. The MICS-EAGLE initiative seeks to enhance the use of data generated by the sixth and latest round of MICS (MICS6) to improve education policies and practices, and thereby achieve better education outcomes for children. The initiative aims to publish data for a total of 70 regions across 61 countries by 2022, creating impressive momentum for further data analysis, especially in the education sector. In fact, several new modules were developed for MICS6, including one that measures the foundational learning skills of children aged 7 to 14 years old, making it the only household survey that collects these data.

Analysis of MICS6 data is critical to better understanding the education sector, particularly in measuring and monitoring progress towards Sustainable Development Goal (SDG) 4. By promoting a more robust analysis of MICS6 data, the MICS-EAGLE initiative aims to equip countries with better tools for advocacy, monitoring and planning. Chapter 2 clarifies the link between MICS6 and SDG4 and highlights the importance of bridging data gaps using household surveys.

By offering guidance that considers various aspects of individual countries' backgrounds, the MICS-EAGLE initiative aims to adapt to country-specific needs.

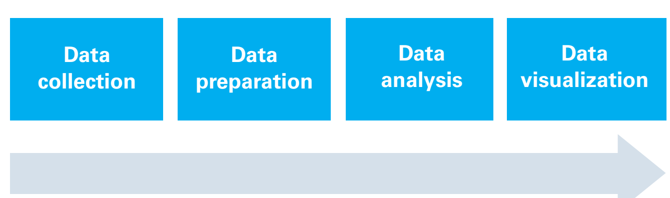
The MICS-EAGLE initiative emphasizes consultative processes with national partners and customizes data analysis to reflect local needs, which are followed by discussion between various stakeholders: local and regional UNICEF offices, national statistical offices and ministries of education.

As part of the MICS-EAGLE initiative, this guide offers an overview of the steps in the data life cycle (Figure 1): data collection and preparation, data analysis, and data visualization. Data preparation, discussed in Chapter 3, includes the preparation of a data set before analysis. Once the data are cleaned, recoded and ready to use, they can be the object of statistical analysis, as presented in Chapters 4, 5 and 6.

The last stage of the data life cycle, data visualization, provides an important link between the data expert and the reader. Understanding data visualization is key to ensuring that an analysis is successfully understood and that the right message is conveyed. All chapters presenting data analysis include a discussion on how to better present the results from the analysis; Chapters 5 and 6 offer an essential overview of several main strategies used to visualize the analytical results of data.

Finally, Chapter 7 introduces a few examples from recent analyses carried out by local governments and UNICEF to better understand country-level education contexts and link data to policy discussion and action. The first two countries to release MICS6 microdata were the Lao People's Democratic Republic (Lao PDR) and Sierra Leone, and they are the focus of the final chapter's analysis.

FIGURE 1 **Data life cycle**



CHAPTER 1

MICS data

For more than 20 years, UNICEF has invested in data collection and helped transform the data landscape. Its international household survey programme, MICS, is the centrepiece of UNICEF's data strategy, allowing the agency to assist countries in collecting and analysing data to fill gaps in monitoring the situation of children and women.

MICS findings have been used extensively as a basis for policy decisions, programme interventions and advocacy to influence public opinion on the situation of children and women around the world.

Government institutions typically carry out the surveys with technical and financial assistance from UNICEF and its partners. UNICEF provides technical support and training to national government staff through a series of regional workshops that cover survey design, data processing, data interpretation, further analysis and dissemination.

MICS data are collected through face-to-face interviews in carefully selected nationally or sub-nationally representative samples of households. High-quality data can be obtained thanks to thorough and tested field procedures, combined with rigorous data verification. After the first results are published, all MICS data sets and Survey Finding Reports can be accessed and downloaded on the programme's website, mics.unicef.org.

MICS history and geographic presence

MICS was originally developed in response to the World Summit for Children as a tool to measure progress towards the Summit's internationally agreed set of mid-decade goals. The first round of MICS was conducted around 1995 in more than 60 countries.

In response to an increased demand for data all over the world, UNICEF has been providing assistance to countries at more frequent intervals. This is providing the opportunity for countries to capture rapid changes in key indicators, as evidenced in the Sustainable Development Goals. As a key source of data on equity, MICS is playing an essential role in tracking progress toward the elimination of disparities and inequities in child well-being. While UNICEF and partners work with national governments to accelerate improvements in the lives of the most vulnerable, MICS can generate the data needed to validate the results of these focused interventions.

Data structure

MICS is typically designed to produce nationally representative estimates. The survey primarily collects data on issues that may affect children's lives. MICS routinely collect household and individual background data so that disparities associated with age, gender, education, wealth, location of residence, ethnicity and other characteristics are revealed with further analysis.

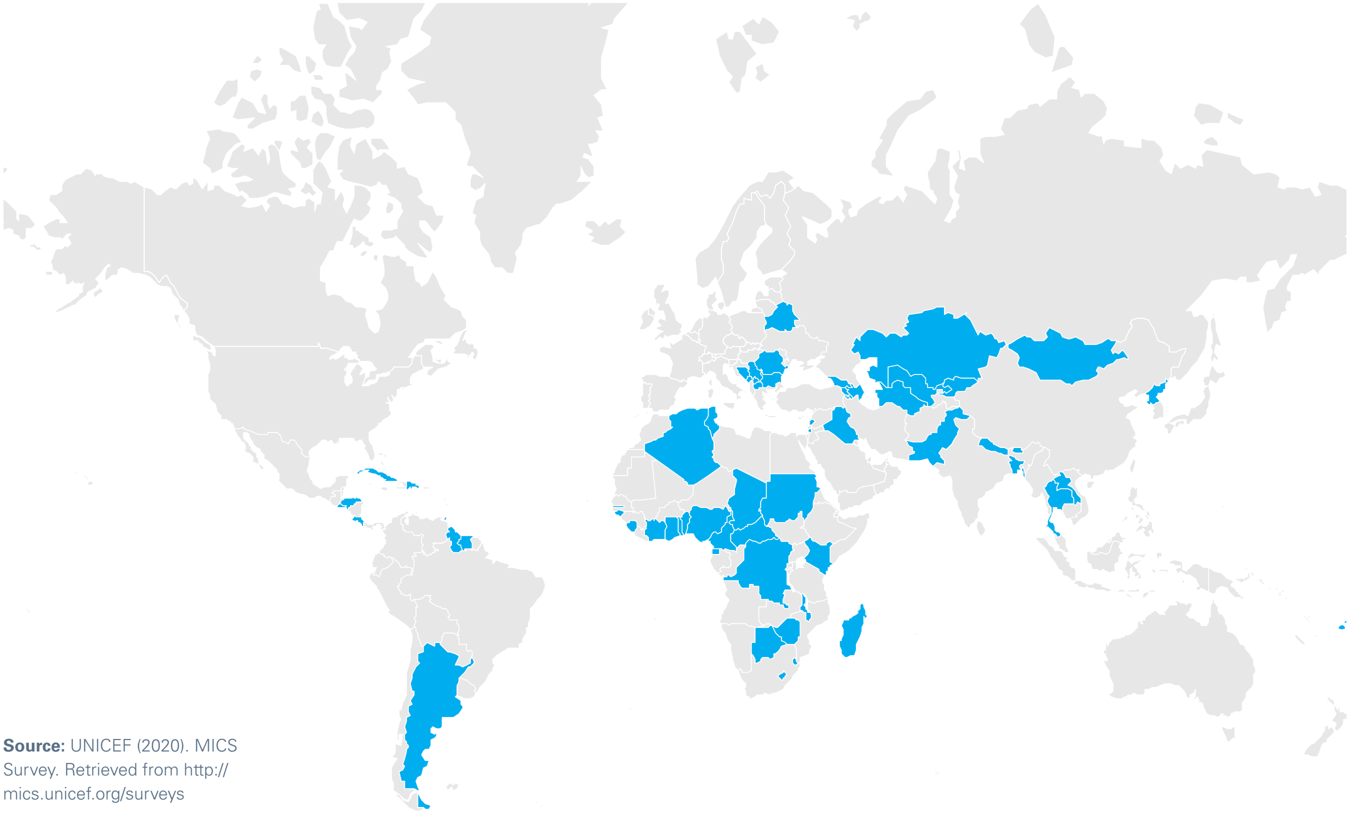
Link with DHS and other surveys

UNICEF works closely with other surveys, such as the Demographic and Health Surveys (DHS) programme, to harmonize methodologies and indicators used in MICS. There is also strong interest from policymakers and researchers in the creation of new modules that can be applied to different surveys. For example, some DHS, such as Senegal DHS 2010–11, include one or more MICS-specific modules, and some MICS include DHS modules.

In recognition of the increasing demand for high-quality household-level survey data as part of the broader global data agenda, the global MICS programme launched a formal partnership with DHS. The three programmes agreed on a structured process to facilitate ongoing collaboration and exploit potential synergies. The goal of the partnership is to increase the frequency, quality and relevance of household survey data around the world and to better serve countries in meeting their domestic and international data demands through improved comparability and integration across surveys, enhanced survey methods and techniques, and greater coordination on survey timing and scheduling.

The techniques for analysing education indicators presented in this manual can also be a valuable resource for those seeking to use DHS or other sources of household data to create sound education policy.

FIGURE 2 **Countries releasing data under MICS6**



Source: UNICEF (2020). MICS Survey. Retrieved from http://mics.unicef.org/surveys

BOX 1: Administrative data and survey data

Education Management Information Systems (EMIS) are the main source of administrative data in the education sector. An EMIS is defined as "a system for the collection, integration, processing, maintenance and dissemination of data and information to support decision-making, policy analysis and formulation, planning, monitoring and management at all levels of an education system." EMIS data often provides information on diverse topics, such as the total number of children enrolled per level of education, teacher salaries, government expenditures on various educational activities, the structure of the education system, and how funds are split between public and private schools.

In contrast, survey data is collected by administering a questionnaire to a sample of respondents that is representative of a given population. Sample surveys are an extremely important tool, as collecting data from an entire population is very costly and administrative data on many issues is often unavailable. This manual focuses on household survey data, and various education indicators relying on this type of information are presented in Chapter 4.

Sometimes, household and administrative data can be used to calculate the same indicators, which can lead to inconsistent results. Conceptually, administrative and census data are not subject to sampling errors the way surveys can be. This means that properly collected administrative sources should be trusted when available. However, administrative data is often insufficient to respond to key policy questions, and in such cases quality survey data must be used.

Questionnaires and modules[1]

MICS6 proposes eight questionnaires, shown in Figure 3, with a variety of modules that are indicated in the figure by their abbreviations. Each country's national steering committee, which is normally comprised of the National Statistics Office, ministries such as education, health, and labour, and international organizations including UNICEF, decides on the final set of questionnaires and modules to be included in their survey based on an assessment of a country's data gaps.

The standard MICS questionnaires, which were designed by UNICEF in close coordination with partners and other international survey programmes, are customized to reflect the local needs of each country with support from UNICEF's MICS experts. All survey activities, from fieldwork to report writing, are carried out by the national counterparts with continuous technical support from UNICEF.

FIGURE 3 MICS6 questionnaires and modules

HOUSEHOLD QUESTIONNAIRE		QUESTIONNAIRE FOR INDIVIDUAL WOMEN AGED 15–49 YEARS		QUESTIONNAIRE FOR INDIVIDUAL MEN AGED 15–49 YEARS		QUESTIONNAIRE FOR CHILDREN AGED 5–17 YEARS FOR ONE RANDOMLY SELECTED CHILD AGED 5–17 YEARS IN EACH HOUSEHOLD		QUESTIONNAIRE FOR CHILDREN UNDER FIVE	
HH	Household Information Panel	**WM**	Woman's Information Panel	**MWM**	Man's Information Panel	**FS**	5–17 Child Information Panel	**UF**	Under-Five Child Information Panel
HL	List of Household Members	**WB**	Woman's Background	**MWB**	Man's Background	**CB**	Child's Background	**UB**	Under-Five's Background
ED	Education [3+]	**MT**	Mass Media and ICT	**MMT**	Mass Media and ICT	**CL**	Child Labour	**BR**	Birth Registration
HC	Household Characteristics	**CM/ BH**	Fertility/Birth History	**MCM**	Fertility	**FCD**	Child Discipline [5–14]	**EC**	Early Childhood Development
ST	Social Transfers	**DB**	Desire for Last Birth	**MDV**	Attitudes Toward Domestic Violence	**FCF**	Child Functioning	**UCD**	Child Discipline [1–4]
EU	Household Energy Use	**MN**	Maternal and Newborn Health	**MVT**	Victimization	**PR**	Parental Involvement [7–14]	**UCF**	Child Functioning [2–4]
TN	Insecticide-Treated Nets	**PN**	Postnatal Health Checks	**MMA**	Marriage/Union	**FL**	Foundational Learning Skills [7–14]	**BD**	Breastfeeding and Dietary Intake [0–2]
WS	Water and Sanitation	**CP**	Contraception	**MAF**	Adult Functioning [18–49]			**IM**	Immunization [0–2]
HW	Handwashing	**UN**	Unmet Need	**MSB**	Sexual Behaviour			**CA**	Care of Illness
SA	Salt Iodization	**FG**	Female Genital Mutilation	**MHA**	HIV/AIDS			**AN**	Anthropometry
		DV	Attitudes Toward Domestic Violence	**MMC**	Circumcision				
		VT	Victimization	**MTA**	Tobacco and Alcohol Use				
		MA	Marriage/Union	**MLS**	Life Satisfaction				
		AF	Adult Functioning [18–49]						
		SB	Sexual Behaviour						
		HA	HIV/AIDS						
		MM	Maternal Mortality						
		TA	Tobacco and Alcohol Use						
		LS	Life Satisfaction						
WQ	WATER QUALITY QUESTIONNAIRE For a subset of households within each cluster							**HF**	QUESTIONNAIRE FOR VACCINATION RECORDS AT HEALTH FACILITY For countries where all the immunizations records are kept in health facilities
GP	GPS DATA COLLECTION QUESTIONNAIRE For countries without existing cluster location data								

Source: UNICEF (2020). MICS Survey. Retrieved from http://mics.unicef.org/surveys

Education-specific data (questionnaires and modules)[2]

MICS6 generates rich education data, but the data are scattered across various parts of the questionnaires. Modules in the household questionnaire, individual men and women questionnaires, the questionnaire for children under 5 and the questionnaire for children aged 5–17 years all contain information that can be used for education policy analysis. Figure 13 in Chapter 4 provides a detailed list of key education indicators and the questionnaires that gather the data used to calculate them.

Household questionnaire

Use of HH questionnaire for education specific analysis: *In terms of education indicators, HH questionnaires serve as the main data source for calculating attendance rates (gross and net), completion rate (SDG4.1.4), out-of-school children (SDG4.1.5), gross intake ratio (SDG4.1.3), effective transition rate, parity indices (SDG4.5.1), repetition rate, dropout rate, over-age students for grade (SDG4.1.6), participation rate in organized learning – one year before the official primary entry age (SDG4.2.2), and school readiness.*

The Education module (ED module) in the household questionnaire hosts the basic sets of education indicators including:

- highest level of education attended by everyone in the household, including adults (ED5);
- completion of education (ED6);
- level of education currently attending (ED10);
- type of education institution being attended (ED11);
- school tuition and other kinds of material support (ED12, ED13 and ED14); and
- level of education attended in the previous school year (ED16).

In addition, the Social Transfer module (ST module) asks about various external economic assistance programmes provided to households. In certain cases, this module can include information on the public and private provision of education.

Men and women questionnaires

Use of men and women questionnaires for education specific analysis: *These questionnaires provide data on critical cross-sectoral analysis such as ICT skills by gender (SDG4.4.1), literacy level by gender (SDG4.6.2), and early marriage by gender (SDG5.3.1). Each of these can be analysed using school attendance or educational attainment to check for correlations with education.*

Men and women questionnaires collect data from men and women aged between 15 and 49 years old. The Women's Background (WB) and Men's Background (MWB) modules provide information on several variables relevant to education. Questions that are also present in the household questionnaire should use consistent wording to ensure the comparability of results. Questions included in the background module that are relevant for calculation of education indicators are:

- completion of education (WB7 for women and MWB7 for men);
- level of education currently attending (WB9 for women and MWB9 for men);
- level of education attended in the previous school year (WB12 for women and MWB12 for men); and
- ability to read a sentence (WB14 for women and MWB14 for men).

The Marriage/Union modules (MA for women and MMA for men) provide information on the age of first marriage, which can be used to estimate child marriage. As early marriage is often a deterrent to education, it is important to include age of first marriage as a disaggregation level for education outcomes such as literacy.

The Mass Media and ICT modules (MT for women and MMT for men) provide information on men's and women's weekly use of information and communication technology (ICT) skills. Men and women are asked which of nine ICT skills they used on a weekly basis over the course of the three months prior to the interview.

FIGURE 4 Activities measured to assess ICT skills

- Copy or move a folder
- Connect and install new devices
- Create electronic presentations with presentation software
- Find, download, install and configure software
- Send e-mails with attached files
- Transfer files between a computer and other devices
- Use basic arithmetic formulae in spreadsheet
- Use copy or paste tools to duplicate or move information within a document
- Write a computer programme using a specialised programming language

Source: UNICEF (2019). *ICT skills divide: Are all of today's youth prepared for the digital economy?* Retrieved from https://blogs.unicef.org/evidence-for-action/ict-skills-divide-to-days-youth-prepared-digital-economy.

BOX 2: Mass Media and ICT module

Digital skills are those that allow children and adolescents to become "digitally literate," able to not only use and understand digital technology, but also to create and share digital content, build knowledge and solve problems. As shown in Figure 4, ICT skills relate to activities that involve information and communication technology.

In order to enable governments to measure the use of ICT skills within their populations, UNICEF updated the Mass Media and ICT module as part of MICS6. The new module, which is adapted from the ITU manual,[3] collects data for men and women aged 15–49 years on their access to and frequency of use of the internet, computers and mobile phones. Moreover, to assess the prevalence of ICT skills, the module collects information on the recent use of ICT skills by measuring certain activities. The ICT-related activities range from less complex to more complex tasks, providing insights into ICT proficiency in participating countries.

A blog post on this topic is available at https://blogs.unicef.org/evidence-for-action/ict-skills-divide-todays-youth-prepared-digital-economy.

Children under 5 years old

Use of children under 5 years old questionnaire for education specific analysis: *This questionnaire provides the data for calculating UNICEF's early childhood development index (ECDI) (SDG4.2.1) and ECE attendance.*

Early learning and development are fundamental for later stages of education. The questionnaire for children under five provides information on:

- early childhood education (ECE) attendance (question UB8).

Information on early schooling can also be connected to data on how children have developed, which is present in the Early Childhood Development module (EC). The module asks questions about parental involvement and the home environment (questions EC1 to EC5), as well as on the child's health and development (EC7 to EC15). The questions on child development are then used to calculate ECDI, which is used to track SDG4.2.1. – Proportion of children under 5 years of age who are developmentally on track in health, learning and psychosocial well-being, by sex.

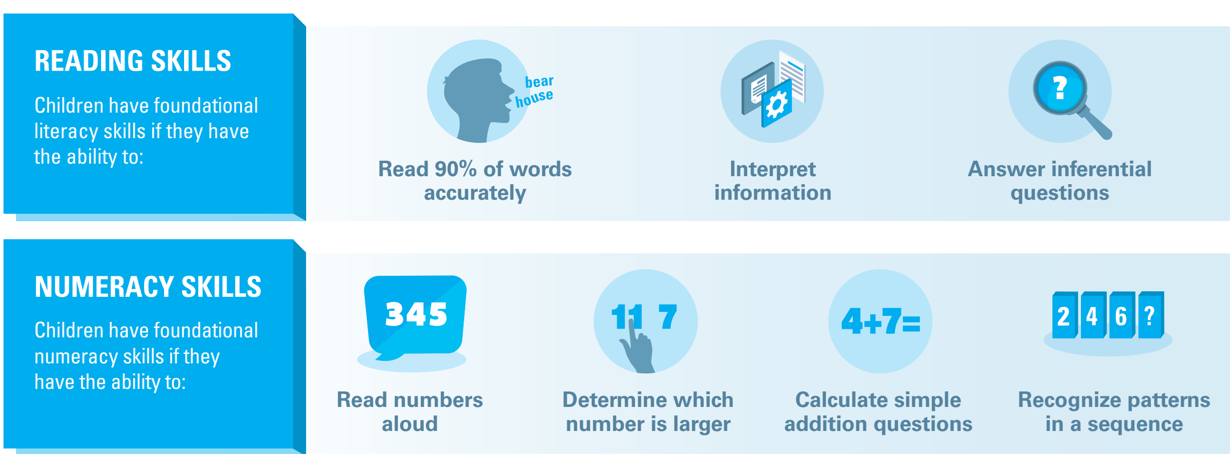
Similar to the questionnaire targeting children aged 5 to 17 years, some modules in the questionnaire for children under 5 provide data to analyse equity in education, as well as determinants of education. The Child Discipline module (UCD) in this questionnaire is analogous to the MICS6 FCD module for children aged 5 to 17 years, but adapted to children aged 2 to 4 years.

BOX 3: Foundational Learning Skills module

This new tool assesses children's reading skills by asking them to read a short story aloud, after which they are asked five questions related to the text. Children are considered to have foundational reading skills if they can successfully read 90 per cent of the words in the text and correctly answer questions related to the story, interpreting and inferring the information contained therein.

In the case of math, children are considered to possess foundational skills when they perform adequately in the following areas: number recognition (the ability to read numbers), number discrimination (the ability to determine which of two numbers is larger), simple addition, and pattern recognition using sequences of numbers.

FIGURE 5 Activities measured in MICS Foundational Learning Skills modules



Source: UNICEF (2019). *How much do children learn? New evidence from Sierra Leone.* Retrieved from https://blogs.unicef.org/evidence-for-action/much-children-learn-new-evidence-sierra-leone.

BOX 4: The Child Functioning module

As access to education increases around the world, so does the concern about schools providing the necessary support to children with disabilities. Too frequently, unaccommodating environments for students with functional difficulties prevent them from making the most of their educational opportunities. In many countries data on these disadvantaged children and how they participate in school lack quality and comparability. To address this gap, UNICEF, in collaboration with the Washington Group on Disability Statistics (WG), has developed a new questionnaire on children with disabilities to be administered in household surveys. The new module collects data on 12 functional domains for children aged 5 to 17 and is included as part of the MICS6 survey.

FIGURE 6 The 12 functional domains measured by the new UNICEF/WG Child Functioning module

ACCEPTING CHANGE
Difficulty accepting change in their routine

AFFECT/EMOTION

ANXIETY
Seeming very anxious, nervous or worried on a daily basis

DEPRESSION
Seeming very sad or depressed on a daily basis

COMMUNICATION
Difficulty being understood by people

CONCENTRATING
Difficulty concentrating on an activity they enjoy doing

CONTROLLING BEHAVIOUR
Difficulty with controlling their behaviour

HEARING
Difficulty hearing sounds like peoples' voices or music

LEARNING
Difficulty learning things

MAKING FRIENDS
Difficulty making friends

REMEMBERING
Difficulty remembering things

SEEING
Difficulty seeing or cannot see at all

SELF-CARE
Difficulty with self-care such as feeding or dressing himself/herself

WALKING
Difficulty walking on level ground

Source: UNICEF (2019). *Do children with disabilities attend school? New findings from Sierra Leone.* Retrieved from https://blogs.unicef.org/evidence-for-action/children-disabilities-attend-school-new-findings-sierra-leone.

Note: Anxiety and depression are both manifestations of the same functional domain: Affect.

Children aged 5–17 years

Use of children aged 5–17 years questionnaire for education specific analysis: *This questionnaire provides the data for calculating learning levels (SDG4.1.1.a) and information on child functioning (SDG4.5), parental involvement, positive and stimulating home environment (SDG4.2.3) and child labour (SDG 5.3.1). It is therefore useful for both education related information and cross-sectoral analysis.*

The questionnaire for children aged 5–17 years contains data on learning in its Foundational Learning Skills module (FL). This module provides valuable information for tracking:

- how much children have learned in reading skills (questions FL10 to FL22); and
- how much children have learned in numeracy skills (questions FL23 to FL27)

This questionnaire also gathers a wealth of data on children's characteristics that can be used to understand disaggregated education attendance, completion and learning by various characteristics. For example, child labour, which can make it more difficult for students to remain in school, can be analysed using the Child Labour module (CL). Assessment of the use of several types of parental disciplinary actions towards their children, which can also impact a child's education, can be found in the Child Discipline module (FCD). Another important module is Child Functioning (FCF), which discusses disabilities that education systems are often not prepared to accommodate. Finally, the Parental Involvement module (PR) asks parents or primary caretakers several questions about how they help and stimulate their children. The PR module can be used to explain the power of parental engagement in guaranteeing successful learning outcomes for children.

CHAPTER 2

SDG4 and international comparative analysis

The SDGs are a collection of 17 global goals for 2030 set by the United Nations General Assembly and adopted by all United Nations Member States in 2015. The goals are part of Resolution 70/1 of the United Nations General Assembly, the 2030 Agenda for Sustainable Development. SDG4 aims to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all."

Indeed, education plays a central theme throughout the 2030 Agenda, which includes a standalone education goal and education-related targets within seven other SDGs.[4]

SDG4 targets and indicators (global and thematic)

Obtaining a quality education is the foundation to creating sustainable development. Access to inclusive education not only improves quality of life, but can also help equip local governments with the tools required to develop innovative solutions to some of the world's most difficult problems. SDG4 has 10 targets, of which seven are expected outcomes and three are means of achieving these targets.

Under these targets, there are 11 global indicators and 32 thematic indicators agreed by the United Nations Statistical Commission and the Technical Cooperation Group to monitor SDG4, as shown in Figure 7. All countries are required to report and monitor the 11 global indicators to the UNESCO Institute for Statistics (UIS), while they can opt to monitor the thematic and other national indicators that fit their needs. Many of these indicators can be calculated using MICS data, a full list of which can be found in the digital annex.[5]

FIGURE 7 Full list of SDG4 education-related indicators (including 13 global indicators and 32 thematic indicators).[6]

Primary and secondary education	Target 4.1 By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes		For monitoring in 2018	Requires further development
Learning	4.1.1	Proportion of children and young people (a) in Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex	YES	YES
	4.1.2	Completion rate (primary education, lower secondary education, upper secondary education)	YES	
Completion	4.1.3	Gross intake ratio to the last grade (primary education, lower secondary education)	YES	
	4.1.4	Out-of-school rate (primary education, lower secondary education, upper secondary education)	YES	
Participation	4.1.5	Percentage of children over-age for grade (primary education, lower secondary education)	YES	
	4.1.6	Administration of a nationally-representative learning assessment (a) in Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education	YES	
Provision	4.1.7	Number of years of (a) free and (b) compulsory primary and secondary education guaranteed in legal frameworks	YES	

Early Childhood	Target 4.2 By 2030, ensure that all girls and boys have access to quality early childhood development, care and preprimary education so that they are ready for primary education		For monitoring in 2018	Requires further development
Readiness for primary school	4.2.1	Proportion of children under 5 years of age who are developmentally on track in health, learning and psychosocial well-being, by sex *Note: measuring this for infants aged 0 to 23 months globally has been recognized as not feasible*	YES	YES
Participation	4.2.2	Participation rate in organized learning (one year before the official primary entry age), by sex	YES	
Readiness for primary school	4.2.3	Percentage of children under 5 years experiencing positive and stimulating home learning environments	NO	YES
Participation	4.2.4	Gross early childhood education enrolment ratio in (a) pre-primary education and (b) early childhood educational development	YES	
Provision	4.2.5	Number of years of (a) free and (b) compulsory pre-primary education guaranteed in legal frameworks	YES	
TVET and Higher Education	**Target 4.3 By 2030, ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university**		**For monitoring in 2018**	**Requires further development**
Participation	4.3.1	Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex	YES	
	4.3.2	Gross enrolment ratio for tertiary education by sex	YES	
	4.3.3	Participation rate in technical-vocational programmes (15- to 24-year-olds) by sex	YES	
Skills for work	**Target 4.4 By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship**		**For monitoring in 2018**	**Requires further development**
Skills	4.4.1	Proportion of youth and adults with information and communications technology (ICT) skills, by type of skill	YES	YES
	4.4.2	Percentage of youth/adults who have achieved at least a minimum level of proficiency in digital literacy skills	NO	YES
	4.4.3	Youth/adult educational attainment rates by age group and level of education	YES	YES to simplify
Equity	**Target 4.5 By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations**		**For monitoring in 2018**	**Requires further development**
Policy	4.5.1	Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated	YES	
	4.5.2	Percentage of students in primary education who have their first or home language as language of instruction	NO	YES
	4.5.3	Extent to which explicit formula-based policies reallocate education resources to disadvantaged populations	NO	YES
	4.5.4	Education expenditure per student by level of education and source of funding	YES	
	4.5.5	Percentage of total aid to education allocated to least developed countries	YES	
Literacy and Numeracy	**Target 4.6 By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy**		**For monitoring in 2018**	**Requires further development**
Skills	4.6.1	Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex	YES	YES
	4.6.2	Youth/adult literacy rate	YES	
Participation	4.6.3	Participation rate of illiterate youth/adults in literacy programmes	YES	

Global Citizenship	Target 4.7 By 2030, ensure that all learners acquire the knowledge and skills needed to promote sustainable development, including, among others, through education for sustainable development and sustainable lifestyles, human rights, gender equality, promotion of a culture of peace and non-violence, global citizenship and appreciation of cultural diversity and of culture's contribution to sustainable development		For monitoring in 2018	Requires further development
Provision	4.7.1	Extent to which (i) global citizenship education and (ii) education for sustainable development are mainstreamed in (a) national education policies; (b) curricula; (c) teacher education; and (d) student assessment	YES	
	4.7.2	Percentage of schools that provide life skills-based HIV and sexuality education	YES	
	4.7.3	Extent to which the framework on the World Programme on Human Rights Education is implemented nationally (as per the UNGA Resolution 59/113)	NO	YES
Knowledge	4.7.4	Percentage of students by age group (or education level) showing adequate understanding of issues relating to global citizenship and sustainability	NO	YES
	4.7.5	Percentage of students in the final grade of lower secondary education showing proficiency in knowledge of environmental science and geoscience	NO	YES
	4.7.6	Extent to which national education policies and education sector plans recognize a breadth of skills that needs to be enhanced in national education systems	NO	YES
School Environment	**Target 4.a Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, nonviolent, inclusive and effective learning environments for all**		**For monitoring in 2018**	**Requires further development**
Resources	4.a.1	Proportion of schools offering basic services, by type of service	YES	YES
Environment	4.a.2	Percentage of students experiencing bullying in the last 12 months	YES	
	4.a.3	Number of attacks on students, personnel and institutions	YES	
Scholarships	**Target 4.b By 2020, substantially expand globally the number of scholarships available to developing countries, in particular least developed countries, small island developing States and African countries, for enrolment in higher education, including vocational training and information and communications technology, technical, engineering and scientific programmes, in developed countries and other developing countries**		**For monitoring in 2018**	**Requires further development**
Numbers	4.b.1	Volume of official development assistance flows for scholarships by sector and type of study	YES	
Teachers	**Target 4.c By 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing States**		**For monitoring in 2018**	**Requires further development**
Trained	4.c.1	Proportion of teachers with the minimum required qualifications, by education level	YES	
	4.c.2	Pupil-trained teacher ratio by education level	YES	
Qualified	4.c.3	Percentage of teachers qualified according to national standards by education level and type of institution	YES	
	4.c.4	Pupil-qualified teacher ratio by education level	YES	
Motivated	4.c.5	Average teacher salary relative to other professions requiring a comparable level of qualification	NO	YES
	4.c.6	Teacher attrition rate by education level	YES	
Supported	4.c.7	Percentage of teachers who received in-service training in the last 12 months by type of training	NO	YES

Source: Retrieved from http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/04/Official-list-of-all-SDG-4-Indications-April-2020.pdf

MICS data and SDG4

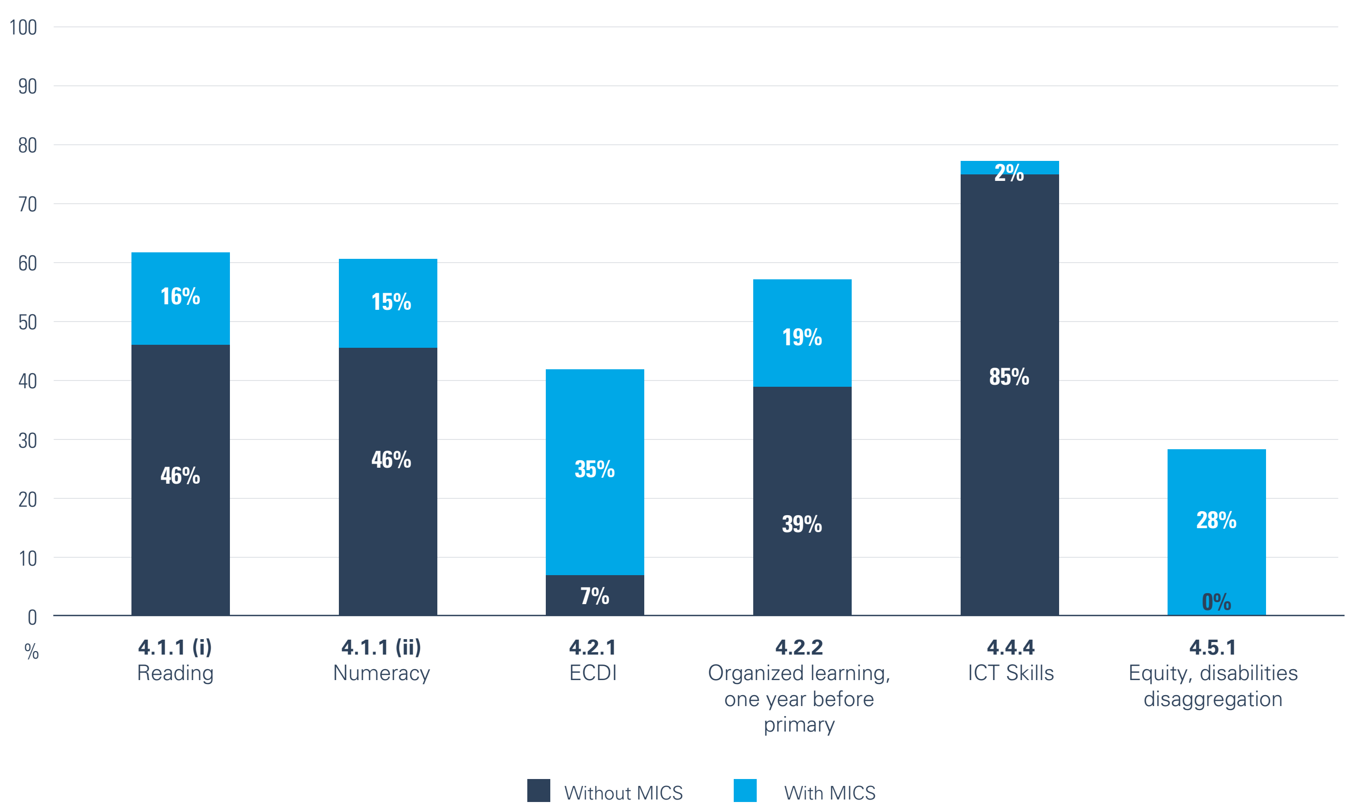
Since its inception in 1995, MICS data have become the largest source of statistically sound and internationally comparable data on women and children worldwide. In countries as diverse as Costa Rica, Mali and Qatar, trained fieldwork teams conduct face-to-face interviews with household members on a variety of topics – focusing mainly on those issues that directly affect the lives of children and women. MICS was a major source of data for the Millennium Development Goal indicators and continues to be a major data source for measuring SDG indicators during the 2030 Agenda.

SDG4 was established in a context where more than half of children and adolescents worldwide are not meeting minimum proficiency standards in reading and mathematics.[7] It serves to refocus efforts on improving access to and quality of education.

Given that SDG4 has a much more ambitious agenda than the Millennium Development Goals, many of the required indicators are newly-developed and data availability is quite low, especially in early grade learning. The recently launched MICS6 has generated a wealth of data about topics that include reading and numeracy skills, ICT skills, disability, and parental involvement. The development of the Foundational Learning module, in particular, has been a cutting-edge initiative to assess children's reading and numeracy skills through household surveys, which includes both children who are in school and out of school.

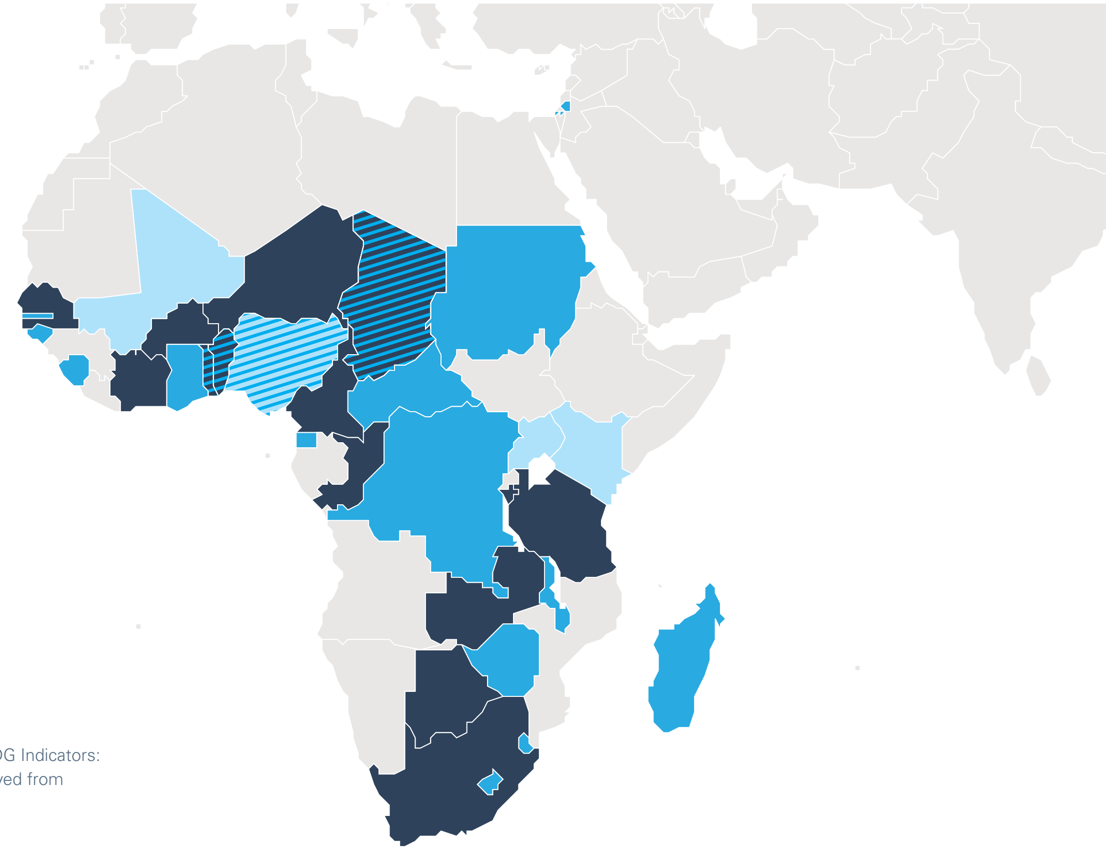
MICS6 has had the highest take-up rate of any MICS round thus far: It is expected that more than 70 regions in over 60 countries will conduct the survey and thus generate new data between 2017 and 2021. Figure 8 presents the contribution of MICS6 to SDG4 data over the next two to three years. These new data will bridge a particularly important gap by providing information on learning in several African countries. As shown in Figure 9, in total, 22 of 49 in countries in sub-Saharan Africa will be covered by MICS6.

FIGURE 8 MICS6 contribution to SDG4 data coverage[8]



Source: UNSTATS (2019). SDG Indicators: Metadata Repository. Retrieved from https://unstats.un.org/home.

FIGURE 9 Data sources for the calculation of foundational learning (SDG4.1.1.a) in sub-Saharan Africa



Source: UNSTATS (2019). SDG Indicators: Metadata Repository. Retrieved from https://unstats.un.org/home.

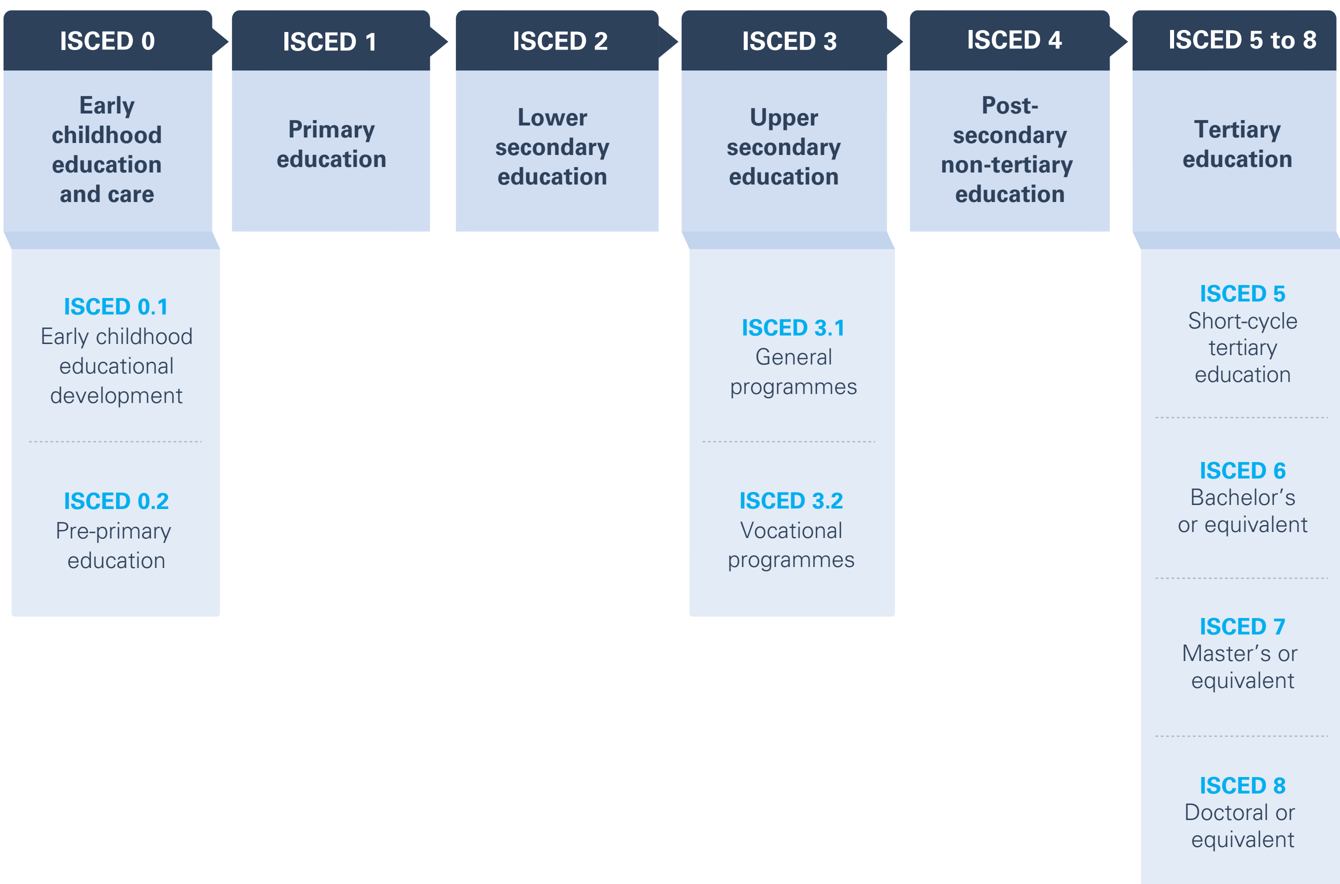
International education data analysis and the ISCED framework

The world's education systems vary widely in terms of structure and curricular content. Consequently, it can be difficult to compare national education systems across countries or to benchmark progress towards national and international goals.

The International Standard Classification of Education (ISCED 2011) is a comprehensive framework used to collect, organize and report education statistics that are internationally comparable. It allows education programmes and academic qualifications to be compared by applying uniform and internationally agreed definitions. A widely used global reference classification for education systems, ISCED is maintained and periodically revised by the UIS in consultation with Member States and other international and regional organizations. ISCED 2011 is the second major revision of this classification (initially developed in the 1970s and revised in 1997), adopted by the UNESCO General Conference in November 2011. The ISCED 2011 Operational Manual[9] provides further guidelines for classifying national education programmes and related qualifications according to ISCED 2011. As seen in Figure 10, ISCED levels can be used to classify all education programmes from early childhood education to those at the doctoral level.

Most education data, including MICS6 data, are collected using national classifications for education levels, which is very important for national policy discussions. However, to be internationally comparable, these data must be re-classified using ISCED. Data processors and users need to pay attention to whether ISCED or a national classification was used in the calculation of a given indicator, as each standard leads to different results.

FIGURE 10 **ISCED 2011 classification framework**



Source: UNESCO-UIS (2012). *International Standard Classification of Education: ISCED 2011*. Paris: UNESCO. http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf.

CHAPTER 3

Collecting, processing and preparing data for analysis

The first step in data analysis is to prepare the data set. This involves various steps to ensure quality, as well as preparation of the data set so it is simpler to analyse. Examples of SPSS and Stata codes used to prepare education data are available in the digital annex.[10]

Data collected by MICS follow strict procedures and concrete processing methods and rules to ensure data quality. This section highlights some of the main tests that should be systematically applied in the different processes of collecting, processing and preparing MICS data.

Data collection

Early rounds of MICS used paper questionnaires, but since its fourth round MICS has also used tablet computers to collect data via a technique called CAPI (Computer Assisted Personal Interviewing), which was generalized to all participating countries and all modules in MICS6. The data collection application used for interviewing improves accuracy by flagging inconsistencies while data collection is ongoing, which enables interviewers to correct the relevant entries.[11]

Nevertheless, inconsistencies across questionnaires persist. For example, an individual's declared level of education on the household questionnaire could be different from that provided on the individual questionnaire. This can happen, for example, when the person responding to the household questionnaire cannot provide accurate information on the level of education attained by each member of the household, whereas an individual's response may be more accurate. Conversely, individual responses can be overestimations. For example, an individual might claim to have attended a higher level of education than they actually did, and another member of the household may be able to give more accurate information. As a result, judgement may be required to determine which of the two data points (self-declared or declared by someone else) provides more accurate information.

Primary data processing

In this stage, national statistical office specialists who have participated in the MICS data processing workshop enter and edit data using software called CSPro, which highlights possible errors in the data. Primary data processing in the MICS programme includes two critical steps:

1. Data entry

A number of consistency and quality checks are done at this stage, regardless of the mode of data collection (paper-based or computer-assisted). Skip patterns defined in the questionnaires are embedded in the data entry/data collection application when tablet-based surveys are used. Data entry operators and interviewers are also alerted to possible inconsistencies in the data. Some inconsistencies must be resolved before moving to the next step of data treatment, while others can remain unresolved and be investigated later in the data editing and quality assurance phase. It is recommended that interviewers resolve all possible inconsistencies in the field – the best time to resolve such issues is while an interviewer is still talking with the respondent.

The main logical checks related to education data run at this stage are to ensure that:

- The highest grade completed at a particular level must be less than or equal to the maximum grade at that level.
- The household member's current level of education (ED10A) cannot exceed his/her highest level of education (ED5A).
- If a household member's current and highest level are the same, his/her current grade of education should not be more than one grade higher than his/her highest grade attended.
- The household member's level of education in terms of attendance in the previous year cannot exceed his/her highest level of education attended.
- If a household member's previous year's and highest level of attendance are the same, his/her previous year's grade of education should not be more than one grade higher than his/her highest grade attended.
- If a household member's previous year's and current level of attendance are the same, his/her previous year's grade of education should not be higher than his/her current grade.
- If a household member's previous year's and current level of attendance are the same, his/her previous year's grade of education should not be lower than his/her current grade minus one.

- If a household member's current year's and highest level of attendance are the same, his/her current year's grade of education should not be the same as his/her highest grade attended, if that grade was reported to be completed.

2. Data editing and quality assurance
The quality of MICS data is further assured by the application of several processes. Upon completion of the data entry, all questionnaires are checked for inconsistencies, which are resolved where possible. This process is strictly defined in the secondary editing guidelines of MICS.[12] Data processing staff are recommended to follow the processes in the guidelines and not make any additional changes.

It is recommended that the quality of MICS data be checked on a regular basis while data collection is ongoing. This is done through the set of quality control tables (field check tables) that are generated every week or two. These tables, which include information on response rates, age displacement, and completeness of data, indicate potential problems in the field. The results of the tables are then passed to the field teams to improve the quality of data, if necessary.

Sample Weights

Sample surveys are normally implemented using a probability-based sampling frame. This means that each individual in the data set represents a number of individuals in the population, and the rate of representation of the individual in the sample to the population is given by the weight. It is important to take the appropriate sampling weight into consideration when calculating all indicators for tables and other descriptive analysis presented in the subsequent sections.

Each questionnaire in MICS represents one part of the population and has its own weighting scheme. As the household questionnaire aims to represent all households in the population, its weighting scheme is used to calculate indicators generalized for the whole population. However, for the other questionnaires, one individual represents a sample of a different population group according to the part of the population the questionnaire covers. For example, when running indicators for the questionnaire for children aged 5–17 years, it is important to use the weights appropriate to that questionnaire.[13]

BOX 5: Preparing your dataset

Before calculating indicators, it is important to prepare the dataset for use. Preparing your dataset includes labelling variables in a way that they can be easily understood. For example, variable ED10A on current level of education attended can be relabelled as "current_edu." Another step in preparing a dataset includes the recoding of variables. For example, in the Laos PDR MICS (Lao Social Indicator Survey II 2017), the variable for location is coded as 1 for urban areas, 2 for rural areas with roads and 3 for rural areas without roads. However, certain analyses will be presented only for rural versus urban areas. In this case, the variable must be recoded as a dummy/binary variable where 1 is urban and 0 is rural.

Merging data sets[14]

In order to calculate some indicators or carry out further statistical analysis, information present in two or more questionnaires must be combined. Those questionnaires must be merged into a single data set. To merge data contained in two questionnaires, it is important to understand what the identifying variables in the data set are. In MICS, every individual is identified by the combination of three variables: HH1 (Cluster number), HH2 (Household number) and HL1 (Line number). The concatenation of the values of these three variables leads to a unique ID for each individual in the data set.

The variables that identify an individual by household number can be used to match that individual across all other questionnaires. For example, if one is interested in disaggregating indicators referring to children under 5 using their household characteristics, it is necessary to merge the data in the household questionnaire with the data in the questionnaire for children under 5. Some characteristics, such as mother's education, are already included by default in the child questionnaire. Others, such as those used in more elaborate analyses that rely on related information, are present in two or more questionnaires and will require merging both data sets.

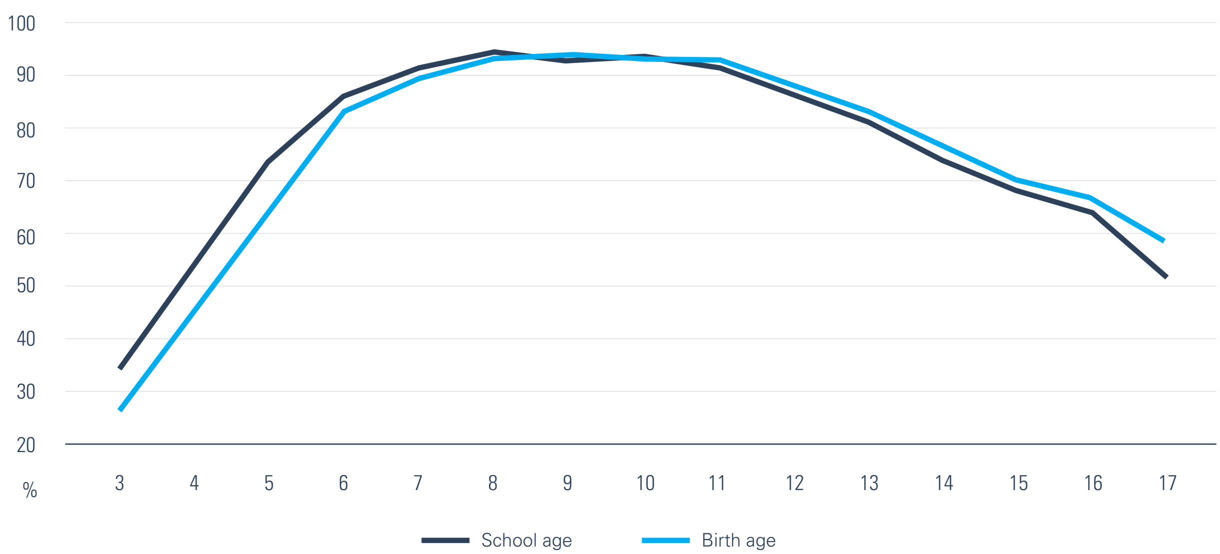
BOX 6: The importance of using the school age variable

Education indicators are strongly impacted by how a child's age is calculated, and different survey methodologies take varying approaches to determining a child's age. Two types of age are commonly used: 1) Birth age (variable HL6 on the household questionnaire of MICS6 surveys), which is a child's chronologic age at the time of the survey; and 2) School age (variable schage in all MICS6 questionnaires), which is defined as a child's age at the beginning of the school year. In MICS surveys, school age is calculated using a child's birth year and month, as well as country-specific start dates for the academic year.

Using the school age variable present in MICS6 surveys increases the quality and comparability of education indicators because it considers two aspects: 1) A child's age may change between the beginning of the school year and the time when a survey interview takes place. By using school age, the age of the child at the beginning of the school year is constant regardless of when the survey takes place; and 2) In most countries, the academic year does not coincide with the calendar year, and may not coincide with other countries' academic years. For example, in Nigeria, the academic year runs from January to December, whereas in Sudan it runs from June to March. That means that children who turn 6 in February of a given year should attend Grade 1 in Nigeria, but Grade 2 in Sudan. By using the beginning of the academic year as common reference, the school age variable guarantees greater comparability across countries.

In the example shown below, most 9-year-olds in Lao PDR are in school, meaning that mistakenly including children who are 8 or excluding children who are 10 will not affect the indicator calculation as much. However, differences in age-specific attendance between ages 3–4 and 16–17 are very strong – at ages 3 and 4, using the school age variable indicates higher attendance than the birth age variable does, whereas the opposite is true for ages 16 and 17. This illustrates that mistakenly including or excluding individuals from these age groups by using birth age instead of school age can lead to significant distortions in indicators.

FIGURE 11 **Age-specific attendance rates in Lao PDR by school age and birth age**



Source: UNICEF (Forthcoming). *Accuracy Matters: Improving Policy Outcomes by using MICS "School Age" to Calculate Education Indicators.* New York: UNICEF.

Missing data and DK (don't know)

Missing data occur when no data value is stored for an observation of a certain variable. This might occur if the respondent refuses to provide certain information or does not know how to answer. Some MICS6 questionnaires include "Don't know" as a possible response, but even when this option is available, interviewers are encouraged to probe for answers and only use "Don't know" when absolutely necessary.

Missingness can take one of two forms: missing at random and missing not at random. Missing at random is very rare and implies that the subsample of the data excluding the observations with missing data carry the same information as the sample including all observations. As a result, under the assumption of complete randomness, the simple exclusion of missing data would not lead to biased estimations.

Most data are not missing at random. Some people are more likely to refuse or be unable to answer certain questions than others. As a result, simply excluding missing data may lead to biased estimates. For example, if less engaged parents are less likely to respond regarding how they participate in their child's education, excluding missing observations would inflate estimates of parental participation.

The same is particularly true for cross-tabulations. In a country with four ethnic groups, for instance, if children from one ethnic group have lower access to education and their parents are less likely to respond because they don't understand the language of the interview, then all statistics by ethnolinguistic group would be biased. In this case, the missing values would occur more often among a given ethnicity, decreasing the representative size of the ethnicity in the survey findings. If the children of parents who fail to respond for linguistic reasons also perform worse in school, the results for that ethnicity would appear better than they really are, because the true count of poorly performing children would not be included in the calculation.

To highlight the presence of such problems, most MICS6 tables will include a line for missing categories. Although the inclusion of missing categories allows for visualizing the total size of the population, it does not correct for non-sample bias. That means that if the missing cases have specific non-random characteristics, the indicator results will still be biased. However, providing an extra line for missing values helps quantify the magnitude of the bias. In the example presented, there would be five lines for disaggregation of major education indicators: One for each ethnicity and one for children with missing information for ethnicity. Figure 12 presents an example from the Lao Social Indicator Survey II Survey Findings Report that demonstrates how data on mother's education included a separate categories for "No information" and "DK/Missing" (in this case, because no caretaker was present). The small amount of data in each did not affect the overall results.

FIGURE 12 **Example of separate presentation of results for "no information" and missing data**

Table LN.2.5: Age for grade

Percentage of children attending primary and lower secondary school who are underage, at age and over-age for grade, Lao PDR, 2017

	Primary school					
	Per cent of children by grade of attendance				Total	Number of children attending primary school
	Under-age	At official age	Over-age by 1 year	Over-age by 2 or more years		
Mother's education	5.9	67.0	11.5	15.5	100.0	4,213
None or ECE	8.4	75.5	8.1	8.1	100.0	5,711
Primary	12.2	81.7	3.4	2.7	100.0	2,006
Lower secondary	15.1	83.9	1.6	1.4	100.0	672
Post secondary/ Non tertiary	14.0	81.1	2.6	2.2	100.0	380
Higher	18.6	79.6	1.0	0.8	100.0	356
No information	(*)	(*)	(*)	(*)	100.0	15
DK/Missing	(*)	(*)	(*)	(*)	100.0	2

Source: Lao Statistics Bureau (2018). *Lao Social Indicator Survey II 2017, Survey Findings Report*. Vientiane, Lao PDR: Lao Statistics Bureau and UNICEF.

CHAPTER 4

Key education indicators and analyses

This section presents various education indicators that can be calculated using MICS6 data. Each indicator is defined in the beginning of the subsection, followed by the necessary calculation methods. The indicators are organized by four topics: completion, internal efficiency, development and skills, and cross-sectoral indicators.

Particular attention should be paid to which survey and which questionnaire data is used as the surveys may collect similar information across different questionnaires. Variable names may also change for each round of MICS; the names used in this manual refer to MICS6. To ensure accurate calculations, it is important to ensure that the data for the indicators are taken from the relevant modules.

The table below provides a synopsis of indicators presented in this guide with their linkage to the SDG4 indicators and the MICS modules and questionnaires.

Each do file/syntax used in the calculation (both for SPSS and STATA) is available in the digital annex[15] and available for download on the MICS-EAGLE website.[16]

FIGURE 13 Summary of indicators

Indicator category	Indicators	MICS questionnaire	MICS module[17]	SDG aligned with
Completion and access to education	Gross attendance ratio (GAR)	Household	ED	4.2.4
	Net attendance rate (NAR)	Household	ED	-
	Adjusted net attendance rate (ANAR)	Household	ED	-
	Completion rate	Household	ED	4.1.4
	Participation in organized learning	Household	ED	4.2.2
	Out-of-school children rate	Household	ED	4.1.5
	Effective transition rate	Household	ED	-
	Gross intake ratio to the last grade	Household	ED	4.1.3
	Parity indices	Household	ED	4.5.1
Internal efficiency	School readiness	Household	ED	-
	Repetition rate	Household	ED	-
	Dropout rates	Household	ED	-
	Percentage of children over-age for grade	Household	ED	4.1.6
Development and skills	Early Child Development Index (ECDI)	Children Under 5 Years	EC	4.2.1
	Foundational learning skills	Children Aged 5–17 Years	FL	4.1.1[18]
	ICT skills	Women and Men	MT for women and MMT for men	4.4.1
	Literacy rate	Women and Men	WB for women and MBM for men	4.6.2 (Youth literacy)
Cross-sectoral indicators impacting education	Positive and stimulating home environment	Children Under 5 Years	UB	4.2.3
	Parental involvement	Children Aged 5–17 Years	PR	-
	Child labour	Children Aged 5–17 Years	CL	8.7.1
	Early marriage	Women and men	MA for women and MMA for men	5.3.1
	Child functioning	Children Aged 5–17 Years	FS	4.5.1 (if parity is calculated)

Participation and completion

This section presents indicators related to participation and completion across the different levels of education. Analysing these indicators can help reveal inequities in participation and completion, which can be used to identify marginalized groups of children and create targeted policies/interventions to address key barriers and issues.

Gross attendance ratio (GAR)

GAR measures the number of students attending a given level of education at any time during the reference academic year, regardless of age, expressed as a percentage of the official school-age population corresponding to the same level of education. It can be divided into three indicators:

- GAR primary – number of children of any age attending primary education divided by the primary school age population
- GAR lower secondary – number of children of any age attending lower secondary of education divided by the lower secondary school age population
- GAR upper secondary – number of children of any age attending upper secondary education divided by the upper secondary school age population

Calculation

GAR is calculated by dividing the number of students attending a given level of education, regardless of age, by the population of the age group that officially corresponds to the given level of education. The following formula is used:

$$GAR_n = \frac{S_n}{P_n}$$

Where:
GAR_n = gross attendance ratio for level n of education
S_n = students attending level n of education
P_n = population aged the official age for level n of education

Net attendance rate (NAR)

NAR measures the percentage of children of a given age group that are attending an education level compatible with their age. It can be divided into three indicators:

- NAR primary – percentage of children of primary school age currently attending primary school
- NAR lower secondary – percentage of children of lower secondary school age currently attending lower secondary school

BOX 7: What's the difference between attendance and enrolment?

Attendance means a child did in fact attend school at a particular point in time, whereas enrolment refers to children listed in school registries. In practice, attendance is typically gathered through household surveys when parents or primary caretakers are asked whether their children attended school in the current week/month/year. In contrast, enrolment is calculated based on school censuses or other administrative sources, which are provided by schools based on the number of children registered in each class. This manual focuses on household surveys and hence gives more emphasis to attendance than enrolment.

- NAR upper secondary – percentage of children of upper secondary school age currently attending upper secondary school

Calculation

For example, NAR for primary level is calculated by dividing the total number of students in the official primary school age range who attended primary education at any time during the reference academic year by the population of the same age group. The following formula is used:

$$NAR_n = \frac{EAP_n}{P_n}$$

Where:
NAR_n = net attendance rate for level n of education
EAP_n = population aged the official age for level n of education attending that level of education
P_n = population aged the official age for level n of education

Adjusted net attendance rate (ANAR)

ANAR measures the percentage of children of a given age that are attending an education level compatible with their age or attending a higher education level. The rate is termed "adjusted" since it includes both groups. It can be divided into three indicators:

- ANAR primary – percentage of children of primary school age currently attending primary or secondary school
- ANAR lower secondary – percentage of children of lower secondary school age currently attending lower secondary school or higher

- ANAR upper secondary – percentage of children of upper secondary school age currently attending upper secondary school or higher

Calculation
For example, the ANAR for primary education is calculated by dividing the total number of students in the official primary school age range who attended primary or secondary education at any time during the reference academic year by the population of the same age group. The following formula is used:

$$ANAR_n = \frac{EAP_n}{P_n}$$

Where:
$ANAR_n$ = adjusted net attendance rate for level n of education
EAP_n = population aged the official age for level n of education attending that level of education or higher
P_n = population aged the official age for level n of education

It is important to note that the ANAR calculation excludes children attending lower levels of education from its numerator. For example, lower secondary school-age children attending primary schools will be counted as in school, but they will not be counted as attending the education level designed for their age group and hence they will be excluded from the ANAR numerator. As a result, in many developing countries, ANAR will underestimate access to education because although in school, many children are attending levels lower than expected for their age group.

Completion rate (SDG4.1.4)

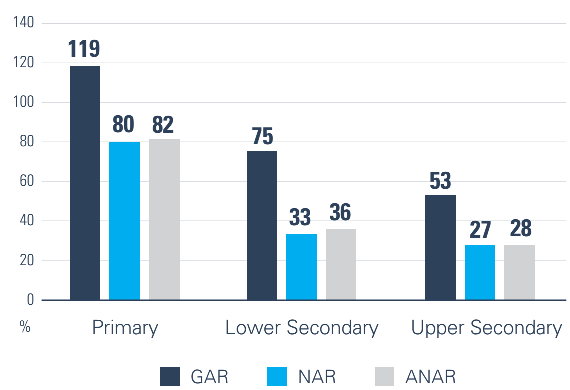
The completion rate reflects the percentage of a cohort of children or young people three to five years older than the intended age for the last grade of each level of education (primary, lower secondary, or upper secondary) who have completed that level of education. The intended age for the last grade of each level of education is the age at which students would enter the grade if they had started school at the official primary entrance age, had studied full-time and had progressed without repeating or skipping a grade. For example, if the official age of entry into primary education is 6 years, and primary school has six grades, then the intended age for the last grade of primary education is 11 years. In this case, the reference age group for calculation of the primary completion rate would be 14-16 years (11 + 3 = 14 and 11 + 5 = 16). This indicator is used to calculate SDG4.1.4 – Completion rate (primary education, lower secondary education, upper secondary education).

BOX 8: Gross attendance ratio, net attendance rate and adjusted net attendance rate

All the three indicators presented previously – GAR, NAR, and ANAR – use the same denominator, which is the number of children of the official age for a given level of education. However, the numerator differs for each of these indicators. For example, GAR is often above 100 per cent in primary education in developing countries because the numerator includes every child attending primary education. In some contexts, children who should be attending secondary education based on their age are still in primary, so they are counted in the numerator, but not in the denominator. In Sierra Leone (2017) for example, the primary GAR is 119 per cent.

When only children attending the level of education designed for their age group are considered in the numerator (NAR), the figures drop sharply. However, NAR does not consider children who are already attending higher levels of education, such as children who are primary school age, but are already in secondary school. To include those children in the numerator, it is necessary to use ANAR, which considers all children attending the level of education designed for them or higher. As a result, ANAR is always higher than NAR, although the difference is usually small.

FIGURE 14 **Differences between GAR, NAR and ANAR in Sierra Leone (2017)**



Source: Statistics Sierra Leone (2018). *Sierra Leone Multiple Indicator Cluster Survey 2017, Survey Findings Report*. Freetown, Sierra Leone: Statistics Sierra Leone.

Calculation
The completion rate is the number of individuals in the relevant age group who have completed the last grade of the given level of education expressed as a percentage of the total population of the same age group. The following formula is used:

$$CR_n = \frac{EAP_{n,a+3to5}}{P_{a+3to5}}$$

Where:
CR_n= completion rate for level n of education
$EAP_{n,a+3to5}$= population aged three to five years above the official entrance age a into the last grade of level n of education who completed level n
P_{a+3to5}= population aged three to five years above the official entrance age a into the last grade of level n of education

Participation in organized learning (SDG4.2.2)

Participation in organized learning measures the share of children one year younger than the official age to start primary school who are attending ECE or primary education. This indicator is essentially the ANAR for one year before primary education and is used to calculate SDG4.2.2 – Participation rate in organized learning (one year before the official primary entry age), by sex.

Calculation
Participation in organized learning is conceptually similar to ANAR. It is calculated as the percentage of children one year younger than the official primary school entry age (as of the beginning of school year) attending ECE or primary school. Both the numerator and denominator include only children aged one year younger than the official entry age for primary school. The following formula is used:

$$PiOL = \frac{E_{AGprim-1}}{P_{AGprim-1}}$$

Where:
$PiOL$= participation in organized learning
$E_{AGprim-1}$= children attending early childhood or primary school aged one year younger than the official entry age for primary school
$P_{AGprim-1}$= children aged one year younger than the official entry age for primary school

Out-of-school children rate (SDG4.1.5)

Out-of-school children are children and young people in the official age range for a given level of education who are not attending either pre-primary,[19] primary, secondary or higher levels of education.

The objective of the out-of-school children rate is to identify the part of the population in the official age range for a given level of education not attending school, in order to formulate targeted policies that can be put in place to ensure they have access to education. It is used to calculate SDG4.1.5 – Out-of-school rate for different levels of education, including primary, lower secondary and upper secondary.

Calculation
The out-of-school children rate is calculated as the share of students of the official age for a given level of education attending pre-primary, primary, secondary or higher levels of education, subtracted from the total population of the same age group.

The following formula is used:

$$OSR_n = \frac{SAP_n - A_{AGn}}{SAP_n}$$

Where:
OSR_n = out-of-school rate for children and young people of the official age for level n of education
SAP_n= population of the official age for level n of education
A_{AGn}= children and young people of the official age for level n of education attending any level of education

When the out-of-school rate is calculated from administrative data, it is typically derived from enrolment instead of attendance.

Effective transition rate

The effective transition rate between levels of education measures the percentage of children who transition to the next level of education without repeating.[20] For example, the effective transition rate from primary to secondary would be the share of students attending first grade of lower secondary education divided by those who were attending the last grade of primary education in the previous year.

Calculation
The effective transition rate is the share of children in the first grade of a school level who completed the last grade of the lower school level in the previous year, divided by

the number of children in the last grade of the lower school level in the previous year who are repeating the last grade of the lower school level in the current year. In short, this indicator aims to measure if students that graduate from one level of education and will attend a higher level of education. The following formula is used:

$$ETR_n = \frac{FG}{LGP}$$

Where:
ETR_n = effective transition rate for level n of education
FG = number of children in the first grade of a school level who completed the last grade of the lower school level in the previous year
LGP = number of children in the last grade of the lower school level in the previous year who are not repeating the last grade of the lower school level in the current year

Gross intake ratio to the last grade (SDG4.1.3)

The gross intake ratio to the last grade of primary school (and similar for lower secondary school) is the ratio of the total number of students currently attending the last grade of primary school for the first time (i.e., who are not repeating the grade) to the total number of children of primary school completion age (age at the beginning of the school year appropriate for the last grade of primary school). It is used to calculate SDG4.1.3 – Gross intake ratio to the last grade (primary education, lower secondary education).

Calculation

The indicator is calculated as the number of children attending the last grade of primary school, minus repeaters, divided by the number of children of primary school completion age at the beginning of the school year. The following formula is used:

$$GIR_n = \frac{A_n - R_n}{BSA_n}$$

Where:
GIR_n = gross intake ratio for level n
A_n = number of children currently attending in the last grade of level n
R_n = children repeating the last grade of level n
BSA_n = children whose age at the beginning of the school year is equal to the age corresponding to the last grade of level n

Parity indices (SDG4.5.1)

Parity ratios are calculated as the ratio of two categories of one indicator. For example, the gender parity index for ANAR is calculated as the ratio of girls' ANAR to boys' ANAR. It can be calculated for any level of education and for various categories: e.g., wealth (ANAR in the poorest quintile divided by ANAR in the richest quintile) or area (ANAR in rural areas divided by ANAR in urban areas).

Additional ratios can be calculated following a similar method (ANAR of the group with the lowest ANAR divided by the group with the highest ANAR). This indicator is used to calculate SDG4.5.1 – Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples, and conflict affected, as data becomes available) for all education indicators.

Normally, the more vulnerable group (e.g., children with disabilities) is used for the numerator, while the other group (e.g., children without disabilities) is used as the denominator. Parity is considered to have been achieved when the parity ratio falls between 0.97 and 1.03.

Calculation

Many indicators can be used to calculate parity indices. A very simple calculation for ANAR parity indices is presented below. Usually, group 1 is presumed to have the lowest ANAR and group 2 to have the highest, but it is not always the case. For gender parity, girls are often treated as group 1 even when their ANAR is higher than boys' ANAR. The following formula is used:

$$PR_{g1/g2} = \frac{ANAR_{g1}}{ANAR_{g2}}$$

Where:
$PR_{g1/g2}$ = parity ratio of group 1 to group 2
$ANAR_{g1}$ = ANAR of group 1
$ANAR_{g2}$ = ANAR of group 2

Internal efficiency

This section presents indicators that impact the internal efficiency of the education system. In simple terms, students entering the education system can be seen as inputs and those graduating can be seen as outputs. School resources including classrooms, teachers and textbooks support this process. Dropout and repetition rates are the most common indicators of internal efficiency. Low dropout and repetition rates indicate high internal efficiency, while high dropout

and repetition rates reveal low internal inefficiency. Furthermore, repetition often leads to a higher share of over-age students in a given grade or a given level of education. As a result, the share of over-aged students is also used to measure the degree of inefficiency of an education system. School readiness is not a measurement of internal efficiency per se. However, strong evidence suggests that students in Grade 1 with pre-primary education tend to learn more, and repeat and dropout less often. Therefore, the school readiness indicator will also be presented at the end of this section.

Repetition rate

The repetition rate measures the share of children in a given grade in a given school year who repeated that grade as a percentage of total number of children who attended the grade in the previous year.

Calculation

The repetition rate is calculated as the number of repeaters in a given grade in a school year divided by the number of children from the same cohort attending the same grade in the previous school year. The following formula is used:

$$RR_g = \frac{R_g}{S_g}$$

Where:
RR_g = repetition rate for grade g
R_g = number of children attending grade g in the current school year who also attended grade g in the previous school year
S_g = number of children who attended grade g in the previous school year

Dropout rate

The dropout rate measures the proportion of children from a cohort attending a given grade in a given school year who are no longer attending school in the following year. It is worth clarifying that children who repeat are still considered to be in school and are therefore not included in the calculation for dropout rate.

Calculation

The dropout rate can be calculated as the share of children who drop out of a grade to the total number of children in that grade in the previous year. The following formula is used:

$$DR_g = \frac{D_g}{S_g}$$

Where:
DR_g = dropout rate for grade g
D_g = number of children who attended grade g in the previous year who are no longer attending school in the current school year
S_g = number of children who attended grade g in the previous school year

Over-age students for grade (SDG4.1.6)

Over-age students for grade are represented by the percentage of students in each grade of a given level of education who are at least two years older than the intended age for that grade. The intended age for a given grade is the age at which students would enter the grade if they had started school at the official entrance age of an education level, had studied full-time, and had progressed without repeating or skipping a grade.

For example, if the official age of entry into primary education is 6 years, children in Grade 4 who are older than 11 are defined as over-age students because the intended age for that grade is 9 years. This indicator is used to calculate SDG4.1.6 – Percentage of children over-age for grade (primary education, lower secondary education)

Calculation

The percentage of over-age students for a grade is calculated by the sum of students in a grade who are two or more years older than the intended age for the grade, divided by total enrolments in the grade. When this indicator is calculated from household survey data, it is typically derived from attendance instead of enrolment. The following formula is used:

$$POAG_g = \frac{OAG_g}{A_g}$$

Where:
$POAG_g$ = percentage of children over-age for grade g
OAG_g = students in grade g who are at least two years older than the intended age for that grade
A_g = total students in grade g

School readiness

School readiness provides information about the share of students in Grade 1 of primary school who have attended some form of ECE. ECE exposes children to a structured learning setting at an early age and prepares them for the transition to primary education. Therefore, children who attend ECE tend to be more prepared for primary school than those who do not.

Calculation

School readiness is calculated as the percentage of children attending Grade 1 of primary school who attended preschool in the previous year. The following formula is used:

$$SR = \frac{A_{ECE}}{A_{prim}}$$

Where:
SR = school readiness in a given year
A_{ECE} = children currently in primary school who attended ECE in the previous year
A_{prim} = children in the first grade of primary school in a given year

BOX 9: What's the difference between preschool, ECCE, ECD and ECE?

Early learning involves various concepts. Preschool, pre-primary education or kindergarten are terms with definitions that vary for each national context, but usually all refer to types of education provided before primary education. According to the ISCED 2011 definition, ISCED 0.2 refers to pre-primary education and covers education programs from age 3 to primary education. In contrast, early childhood care and education (ECCE or ECEC) is more than just preparation for primary school. It seeks to nurture caring, capable and responsible future citizens, and focuses on the holistic development of a child's social, emotional, cognitive and physical wellbeing in order to build a solid and broad foundation for lifelong learning. Also known as early childhood development (ECD), in the ISCED 2011 definition it is classified as ISCED 0.1 and is targeted toward children younger than three years of age.

Finally, early childhood education (ECE) is a broader term that focuses on a child's learning between birth and the age of eight, which corresponds to the early stages of primary education in most countries.

Blog posts on this topic are available at https://blogs.unicef.org/evidence-for-action/better-ways-measure-promote-early-education-lessons-lao-pdr and https://blogs.unicef.org/evidence-for-action/much-children-learn-new-evidence-sierra-leone.

Source: UNESCO-UIS (2012). *International Standard Classification of Education: ISCED 2011.* Paris: UNESCO. Retrieved from http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf.

Development and skills

Quality education is central to the SDG4 agenda, which has expanded its focus from equitable access to education to equitable learning. Developing an evidence base to understand whether children are learning in their classrooms and the skills they are able to develop is a core component of the SDGs. This can be measured in various ways. This section lists indicators that provide information on learning outcomes of children across different levels of education. These indicators can be used to track whether education systems are preparing their children for full social and economic participation and to better understand areas where challenges persist.

Early Child Development Index (ECDI) (SDG4.2.1)

Early childhood development is multidimensional, encompassing several aspects of a child's well-being: physical, social, emotional and mental. UNICEF developed the ECDI to measure the percentage of children under 5 years of age who are developmentally on track in literacy-numeracy, physical, social-emotional, and learning domains. These data can be captured by SDG4.2.1. For each domain, developmentally on track is defined as follows:

- **Literacy-numeracy:** A child is developmentally on track in literacy-numeracy if s/he can do at least two of the following: identify/name at least 10 letters of the alphabet; read at least four simple, common words; and/or know the name and recognize the symbols of all numbers from 1 to 10.
- **Physical:** A child is developmentally on track in this domain if one of the following is true: the child can pick up small objects with two fingers or is generally well enough to play.
- **Social-emotional:** A child is considered developmentally on track in social-emotional development if two of the following are true: the child gets along well with other children; the child does not kick, bite or hit other children; and the child does not get distracted easily.
- **Learning:** A child is developmentally on track in the learning domain if one of the following is true: the child follows simple directions on how to do something correctly and/or when given something to do and is able to do it independently.

The ECDI is a composite index that considers children to be developmentally on track if they meet the criteria for at least three of the four domains.

Calculation

ECDI can be calculated using the MICS module on early childhood development. In MICS6, responses to questions EC6-EC15 are used to determine whether children are developmentally on track in four domains. Consistent with the explanation given above, the variables in standard MICS6 data used to assess whether a child is developmentally on track are shown below. Terms such as EC6 and EC7 represent variables in the questionnaire for children under 5 in the MICS6 microdata set.

- Literacy-numeracy (LN) – this domain is coded as 1 when at least two of the following are present, otherwise it is coded as 0:
  - o EC6=1 represents a child being able to identify/name at least 10 letters of the alphabet
  - o EC7=1 represents a child being able to read at least four simple, common words
  - o EC8=1 represents a child being successful in knowing the names and recognizing the symbol of all numbers from 1 to 10

- Physical (P) – this domain is coded as 1 when at least one of the following is present, otherwise it is coded as 0:
  - o EC9=1 represents a child being able to pick up a small object with two fingers, like a stick or a rock from the ground
  - o EC10=2 represents a child being healthy enough to play

- Social-emotional (SE) – this domain is coded as 1 when at least two of the following are true, otherwise it is coded as 0:
  - o EC13=1 represents a child being able to get along well with other children
  - o EC14=2 represents a child not kicking, biting, or hitting other children
  - o EC15=2 represents a child not getting distracted easily

- Learning (L) – this domain is coded as 1 when at least one of the following is true, otherwise it is coded as 0:
  - o EC11=1 represents a child being able to follow simple directions on how to do something correctly
  - o EC12=1 represents a child being able to do a task independently

The ECDI is calculated as an overall index score representing the percentage of children aged 36 to 59 months who are developmentally on track in at least three of these four domains. The sampled child is considered to be developmentally on track and well-prepared for starting primary school in the areas of health, learning and psychosocial well-being when ECDI is equal to 1. The following formula is used to code the ECDI:

$$ECDI=1 \text{ } if \text{ } LN + \mathrm{P} + SE + L \geq 3, \text{ otherwise } 0.$$

Where:
$ECDI$ represents the Early Childhood Development Index

LN is a binary variable where 1 represents children developmentally on track in the literacy-numeracy domain, calculated using EC6, EC7 and EC8

P is a binary variable where 1 represents children developmentally on track in the physical domain, calculated using EC9 and EC10

SE is a binary variable where 1 represents children developmentally on track in the social-emotional domain, calculated using EC13, EC14 and EC15

L is a binary variable where 1 represents children developmentally on track in the learning domain, calculated using EC11 and EC12

Once the ECDI is obtained, it can then be used to calculate the percentage of children who are developmentally on track:

$$\mathrm{SDT}_{3to4} = \frac{\mathrm{ECDI}_{3to4}}{\mathrm{T}_{3to4}}$$

Where:
SDT_{3to4}= share of children aged 3 to 4 years old developmentally on track
ECDI_{3to4}= children aged 3 to 4 years old who have ECDI equal to 1 according to the formula above
T_{3to4}= total number of children aged 3 to 4 years old

Foundational learning skills (SDG4.1.1.a)

Learning outcomes are key indicators of quality education as they provide evidence on whether schools are equipping children with the foundational skills needed for success. The MICS module on foundational learning skills (FL)[21] in the questionnaire for children aged 5–17 years can be used to measure learning outcomes expected for Grades 2 and 3 in numeracy and reading. This data can be used to calculate SDG4.1.1.a to measure the proportion of children in Grade 2/3 achieving minimum proficiency in (i) reading and (ii) mathematics, by sex. As the FL module is administered to children aged 7–14 years, it can also be used to understand the share of children in that age group who have achieved the minimum proficiency of a student in Grade 2/3.

Children are asked a series of questions to measure whether they are achieving minimum foundational skills in reading and numeracy in Grade 2/3. Foundational reading skills are divided into three categories: 1) word recognition (correctly reading 90 per cent of words in a story), 2) literal questions (replying correctly to all three literal questions), and 3) inferential (replying correctly to both of two inferential questions). If the child succeeds in all three tasks, s/he is considered to have foundational reading skills.

Foundational numeracy skills are divided into four tasks: 1) number reading, 2) number discrimination, 3) addition, and 4) pattern recognition. Each task is composed of several questions and the child must correctly answer all questions to successfully complete the task. If the child succeeds in all four tasks, s/he is considered to have foundational numeracy skills.

Calculation

Foundational reading skills
Applying the explanation above, the variables used to assess children's foundational reading skills are:

- FL19W is a set of numbered questions (FL19W1, FL19W2, FL19W3…) that allows for the calculation of the number of words read correctly in a reading exercise. For example, if a child correctly reads 6 words, then FL19W6 will be equal to 1, otherwise it will be blank or equal to zero.
- FL22A, FL22B, and FL22C are binary variables which represent the three literal questions to indicate if the child successfully answered each of the questions.
- FL22D and FL22E are binary variables which represent two inferential questions to indicate if the child successfully answered each of the two inferential questions.

The following calculation is used:
readsk=1 if (*read_corr*=1) & (*alit*=1) & (*alnfe*=1), otherwise 0.

Where:
readsk represents a binary variable for foundational reading skills, equalling 1 when a child has foundational reading skills and 0 otherwise.

read_corr is a binary variable where 1 represents success in reading 90 per cent of the words in a story correctly, calculated using FL19W.

alit is a binary variable where 1 represents success in responding correctly to three literal questions, calculated using FL22A, FL22B and FL22C

alnfe is a binary variable where 1 represents success in responding correctly to two inferential questions, calculated using FL22D and FL22E

$$FRS_{ng} = \frac{readsk_{ng}}{T_{ng}}$$

Where:
FRS_{ng} = share of children aged n and attending a school grade g who have foundational reading skills
$readsk_{ng}$ = children aged n and attending a school grade g who have *readsk* equal to 1 according to the formula above
T_{ng} = total number of children aged n and attending school grade g

It is worth noting that the indicator can be calculated for any age group or for all age groups (data is collected for children aged 7 to 14 years old). Furthermore, the indicator can be calculated for children of a given age group (or all age groups) who are attending a particular grade.

Foundational numeracy skills
Applying the explanation above, the variables used to assess children's foundational numeracy skills are as follows. Each domain of numeracy skills has multiple questions which are listed with capital letters (e.g., A, B, C):

- FL23 (A, B, C, D, E, F) = 1 if the child correctly responded to a number reading question
- FL24 (A, B, C, D, E) = 1 if the child correctly responded to a number discrimination question
- FL25 (A, B, C, D, E) = 1 if the child correctly responded to addition questions
- FL27 (A, B, C, D, E) = 1 if the child correctly responded to pattern recognition questions

The following calculation is used:
numbskill=1 if (number_read=1) & (number_ dis=1) & (number_add=1) & (number_patt=1), otherwise 0.

Where:
numbskill shows children with foundational numeracy skills

number_read is a binary variable, where 1 represents children correctly answering all the number reading questions, calculated using FL23 (A, B, C, D, E, F)

number_dis is a binary variable, where 1 represents correctly answering all the number discrimination questions, calculated using FL24 (A, B, C, D, E)

number_add is a binary variable, where 1 represents correctly answering to all the addition questions, calculated using FL25 (A, B, C, D, E)

number_patt is a binary variable, where 1 represents correctly answering to all the number pattern tasks, calculated using FL27 (A, B, C, D, E)

$$FNS_{ng} = \frac{numbskill_{ng}}{T_{ng}}$$

Where:
FNS_{ng} = the share of children aged n and attending school grade g who have foundational numeracy skills
$numbskill_{ng}$ = children aged n and attending school grade g who have *numbskill* equal to 1 according to the formula above
T_{ng} = total number of children aged n and attending school grade g

It is worth noting that the indicator can be calculated for any age group or for all age groups (data is collected for children aged 7 to 14 years old). Furthermore, the indicator can be calculated for children of a given age group (or all age groups) who are attending a particular grade.

ICT skills (SDG4.4.1)

MICS6 includes the ICT and Mass Media module, which corresponds to SDG4.4.1 and supports countries in assessing the prevalence of ICT skills among men and women. SDG4.4.1 measures the proportion of youth/adults with ICT skills, by type of skill. For this indicator the specific demographic group used is adults aged 15 to 24 years old.

The data is collected through the questionnaire for individual women (under the variables starting with MT6) and individual men (under the variables starting with MMT6) and can be further disaggregated based on location, level of education, age group, language and wealth.

The module measures the proportion of youth and adults who used at least one of nine ICT skills in the three months leading up to the survey:

- MT6A/MMT6A: Copied or moved a file or folder
- MT6B/MMT6B: Used a copy and paste tool to duplicate or move information within a document
- MT6C/MMT6C: Sent e-mail with an attached file, such as a document, picture or video
- MT6D/MMT6D: Used a basic arithmetic formula in a spreadsheet
- MT6E/MMT6E: Connected and installed a new device, such as a modem, camera or printer
- MT6F/MMT6F: Found, downloaded, installed and configured software
- MT6G/MMT6G: Created an electronic presentation with presentation software, including text, images, sound, video or charts
- MT6H/MMT6H: Transferred a file between a computer and other device
- MT6I/MMT6I: Wrote a computer programme in any programming language

Calculation

ICT skills are calculated by the simple ratio of the number of individuals in a demographic group who used a certain ICT skill, divided by the total number of people in that demographic group. It can be calculated for a certain age or gender. The following formula is used:

$$\mathrm{ICT}_{s,d} = \frac{\mathrm{D}_{s,d}}{\mathrm{P}_{s,d}}$$

Where:
$\mathrm{ICT}_{s,d}$ = share of individuals in a specific demographic group d who possess ICT skill S
$\mathrm{D}_{s,d}$ = number of individuals in a specific demographic group d who used a certain ICT skill S
$\mathrm{P}_{s,d}$ = total number of people in that specific demographic group

Literacy rate (SDG4.6.2)

The literacy rate measures the share of population that can both read and write a short, simple statement about their everyday life. Two common literacy rates are measured:

- Adult literacy rate: for individuals aged 15 years and above
- Youth literacy rate: for individuals aged 15–24 years

Literacy rate measures the basic literacy skills that primary education should equip the population with. Therefore, it is used as an indicator to measure the effectiveness of primary education. Analysing literacy rates by measuring the absolute number of illiterates, over a period of time, helps create a comprehensive picture of the prevalence of basic skills in the population. The literacy rate can be disaggregated by sex, age and location.

Calculation

The literacy rate is calculated by dividing the total number of literate individuals in an age group by the population of the age group. The following formula is used:

$$\mathrm{LR}_a = \frac{\mathrm{L}_a}{\mathrm{P}_a}$$

Where:
LR_a = literacy rate for population in age group a
L_a = number of literate individuals in age group a
P_a= population in age group a

Cross-sectoral indicators

Although schools are extremely important in understanding the effectiveness of education systems, factors external to schools may also affect children's learning and participation. The indicators listed in this section explore these factors.

Positive and stimulating home environment for young children (SDG4.2.3)

Home environments can significantly impact children's learning, school performance and participation in classrooms. MICS data can be used to understand the type of home environment in which children live, thereby measuring SDG4.2.3. A home environment that supports young children by providing them with opportunities to expand and apply their skills and knowledge can be said to be positive and stimulating. The data are collected through the children under 5 questionnaire. To measure a positive and stimulating home environment, MICS collects data on the engagement of parents or other adult household members in a wide range of activities, including reading or looking at picture books; telling stories; singing songs; taking children outside the home; playing; and naming, counting and/or drawing. MICS6 also collects information on the availability of books and toys and provisions for adequate supervision to create a full picture of the home environment.

Calculation

A positive and stimulating environment is measured by the engagement of adult members of the household in activities that contribute to child development and can be measured using variables under EC5 in MICS6.

In MICS6, a child is said to receive adequate support if an adult member of the household engaged in at least four of the six activities presented under EC5 in the three days preceding the interview. Similarly, inadequate support is indicated by engagement in fewer than four activities. The following formula is used:

CS = 1 if EC5A + EC5B + EC5C + EC5D + EC5E + EC5F ≥ 4, otherwise 0.

Where:
CS represents an adult engaging in at least four activities
EC5A is a binary variable, where 1 represents an adult member of the household and the child engaging in reading books or looking at picture books together
EC5B is a binary variable, where 1 represents an adult member of the household and the child engaging in storytelling
EC5C is a binary variable, where 1 represents an adult member of the household and the child singing songs and lullabies together
EC5D is a binary variable, where 1 represents the child being taken out of the home by an adult member of the household
EC5E is a binary variable, where 1 represents an adult member of the household engaging in playful activities with the child
EC5F is a binary variable, where 1 represents an adult member of the household and the child playing, drawing and naming things together

$$PSHE_n = \frac{CS_n}{C_n}$$

Where:
$PSHE_n$ = share of children aged n who live in a positive and stimulating home environment
CS_n = children age n with whom an adult engaged in at least four activities
C_n = children aged n

Parental involvement in school

Parents can be involved in their children's education through learning activities at home, participating in school meetings, tracking their children's development in school and other activities. MICS6 collects information related to parental involvement in primary and secondary education using the questionnaire for children aged 5–17 years.

Calculation

The following binary variables measure involvement by adults in school management in the last year:

- School has a governing body open to parents (PR7)
- Attended meeting called by governing body (PR8)
- Attended meeting to discuss key education issues (PR9A)
- Attended meeting to discuss key financial issues (PR9B)
- Children received a report card (PR10)
- Attended a school celebration or sporting event (PR11A)
- Met with teachers to discuss child's progress (PR11B)

Each of these variables shows one dimension of parental involvement and will yield different results after analysis. Usually, the variable will be presented as a ratio, for example, the share of students whose school has a governing body or the share of students whose parents attended a meeting called by the governing body. It is important in those calculations to exclude children who are not attending school, as parental involvement in school is not applicable. For children not attending school, the PR variables related to school engagement are considered missing. The following formula is used:

$$PI_x = \frac{PRX{=}1}{(PRX{=}0\ ){+}(PRX{=}1)}$$

Where:
PI_x represents share of students whose parents engaged in activity X. X can be chosen from: PR7, PR8, PR9, PR10, PR11A and PR11B
PRX=1 represents students whose parents participated in activity X
PRX=0 represents students whose parents did not engage in activity X

Child labour (SDG8.7.1)

SDG8.7.1 – Proportion and number of children aged 5–17 years engaged in child labour, by sex and age, measures the prevalence of child labour, which is also very important in the context of education. Various education variables can be disaggregated by children working under different conditions. For example, ANAR can be calculated for children involved in child labour and compared to the overall ratio.

The child labour module is administered to all children age 5 to 17 using the questionnaire for children aged 5–17 years. Data are collected for three categories of child labour: economic activities, household chores and hazardous conditions.

Calculation

Children are considered to be in child labour if they engage in at least one of three categories: economic activities, household chores and hazardous conditions. These three categories are coded with the following variables:

- **Economic activities** – variables CL1A, CL1B, CL1C and CL1X describe several economic activities such as working on a plot or farm (CL1A), helping the family business (CL1B) and producing handicrafts (CL1C). Children are considered to be working in economic activities if they respond "yes" to any of these questions.
- **Household chores** – variables CL7, CL9, CL11A, CL11B, CL11C, CL11D, CL11E, CL11F and CL11X describe several household chores such as going shopping for the family (CL11A), cooking (CL11B), and washing clothes (CL11C). Children are considered to be performing household chores if they respond "yes" to any of these questions.
- **Hazardous conditions** – variables CL4, CL5, CL6A, CL6B, CL6C, CL6D, CL6E and CL6X describe hazardous conditions such as carrying heavy loads (CL4), working with dangerous tools (CL5), and working at heights (CL6D). Children are considered to be working under hazardous conditions if they respond "yes" to any of these questions.

Engaging in an economic activity or performing household chores does not automatically suggest that a child is engaged in child labour. To qualify as child labour, a threshold number of hours of economic activity or household chores must be met.

In MICS6, question CL3 asks the number of hours engaged in economic activities. This threshold changes for different age groups. For economic activity, the threshold is established by variable CL3 as follows: A child is coded as engaged in child labour if the number of hours s/he is engaged in economic activities is greater than the threshold:

- Age 5–11: 1 hour or more
- Age 12–14: 14 hours or more
- Age 15–17: 43 hours or more

For household chores, two age-specific thresholds are established by summing the total number of hours engaged in chores like fetching water for household use or collecting firewood for household use, using data from CL8, CL10 and CL12:

- Age 5–14: 28 hours or more
- Age 15–17: 43 hours or more

Working under hazardous conditions, regardless of the number of hours, automatically qualifies as child labour according to some definitions (see box below). As a result, the final indicator for child labour under those definitions includes all children working in economic activities or household chores beyond the age-specific limit on hours, or working under hazardous conditions regardless of the number of hours. The following formula is used:

$$CL = \frac{L_{5-17}}{P_{5-17}}$$

Where:
CL= Share of children aged 5–17 engaged in child labor
L_{5-17}= children aged 5–17 years that were involved in economic activities or household chores above the age-specific thresholds or were working under hazardous conditions during the week preceding the survey
P_{5-17}= population of children aged 5–17 years

BOX 10: Defining child labour

The definition of child labour varies by context and even across UN agencies. The MICS indicator PR.3 includes children working in hazardous activities as child labour. However, to ensure comparability of estimates, in 2018 UNICEF and the ILO decided to exclude engagement in hazardous occupations or working under hazardous conditions from estimates of child labour for the purpose of reporting on SDG8.7.1, which should be used by countries when they report to UNSD. The exclusion of hazardous conditions in reporting is fundamentally due to the need for further methodological work to validate questions aimed at identifying children engaged in hazardous activities.

Source: UNSTATS (2019). *Indicator 8.7.1 - E-Handbook on SDG Indicators.* UN Statistics Wiki. Retrieved from https://unstats.un.org/wiki/display/SDGeHandbook/Indicator+8.7.1.

Early marriage (SDG5.3.1)

Data from MICS can be used to calculate the share of women (or men) married before a certain age, usually 15 or 18, in order to understand the prevalence of early marriage in a country. This indicator can be calculated using the adult questionnaires, which means that the information is only available for men and women aged 15–49 years. Early marriage can be disaggregated into cohorts to understand the impact it has on younger individuals. For example, the indicator can be calculated as the share of women aged 20–24 years who married before age 15 or 18, which is SDG5.3.1. This can be used to further analyse academic performance and progression as well as its impact on society.

Calculation

The general formula for calculating SDG5.3.1 is the percentage of women or men from a certain age group (for example aged 15–49 years or 20–24 years) who first married or entered a marital union before their 15th or 18th birthday. The example below calculates early marriage for women aged 20–24 years, using 15 as the cut-off age for a union to be considered early marriage.

$$EM_{20to24} = \frac{W_{m15\ (20to24)}}{W_{20to24}}$$

Where:
EM_{20to24} = early marriage prevalence among women aged 20–24 years
$W_{m15\ (20to24)}$ = women aged 20–24 years who married before they turned 15
W_{20to24} represents all women aged 20–24 years

In MICS6, the following variables are needed to calculate early marriage:

- For women, variable MA1 asks if they are currently married or living with a partner. If they are currently married, the variable is coded as 1, if they are living with a partner it is coded as 2, and if they are not in a union it is coded as 3. Variable MA5 asks about previous unions and is coded similarly to MA1: if MA5 equals 1 it indicates that the woman has been previously married, if it equals 2 it means that the woman has previously lived together with someone, and if it equals 3 it means that the woman has never been married or lived together with someone. (for men, the respective variables used are MMA1 and MMA5).
- For women, variable MA11 provides the age at first marriage or entry into marital union (for men the respective variable is MMA11).

Child functioning (SDG4.5)

Data on child functioning is collected for all children under 18 through either the questionnaire for children under 5 or the questionnaire for children aged 5–17 years. In the case of children under 5, data on functional difficulties are collected on the following functional domains: seeing, hearing, walking, fine motor, communication, learning, playing, and controlling behaviour. For children aged 5–17 years, data on functional difficulties are collected on the following functional domains: seeing, hearing, walking, self-care, communication, learning, remembering, concentrating, accepting change, controlling behaviour, making friends, and affect (or children with difficulties controlling their emotions, which is calculated using metrics for anxiety and depression).

Calculation

Children are considered to have functional difficulties if they have difficulty in at least one functional domain. For each question on functional domains, values of 3 and 4 mean, respectively, "having a lot of difficulty" in the functional domain or "not being able to perform the function all," for example not seeing at all, or not hearing at all.

For children under 5 years of age, functional difficulty in the individual domains is calculated as follows:

- **Seeing** (UCF7 = 3 or 4)
- **Hearing** (UCF9 = 3 or 4)
- **Walking** (UCF11 = 3 or 4, OR UCF12 = 3 or 4, OR UCF13 = 3 or 4) – includes multiple questions on walking with or without equipment
- **Fine motor** (UCF14 = 3 or 4)
- **Communication** – must satisfy one of the two conditions below:
    - **Understanding** (UCF15 = 3 or 4)
    - **Being understood** (UCF16 = 3 or 4)
- **Learning** (UCF17 = 3 or 4)
- **Playing** (UCF18 = 3 or 4)
- **Controlling behaviour** (UCF19 = 5). UCF19 asks mothers and primary caretakers how likely children are to kick, hit or bite other children. Children are considered to have functional difficulties in controlling behaviour if they are very likely to kick, hit or bite other children

A child under 5 years-old is considered to have functional difficulties if any of the above variables is present. The following formula is used:

CD=1 if S=1 or H=1 or W=1 or FN=1 or CM=1 or L=1 or P=1 or CT=1.

Where:
CD represents children with functioning disability:

- S=1 means the child has functional difficulties associated with seeing
- H=1 means the child has functional difficulties associated with hearing
- W=1 means the child has functional difficulties associated with walking
- FN=1 means the child has functional difficulties associated with fine motor capacity
- CM=1 means the child has functional difficulties associated with communicating
- L=1 means the child has functional difficulties associated with learning
- P=1 means the child has functional difficulties associated with playing
- CT=1 means the child has functional difficulties associated with controlling their behaviour

For children aged 5–17 years, functional difficulty in the individual domains is calculated as follows:

- **Seeing** (FCF6 = 3 or 4)
- **Hearing** (FCF8 = 3 or 4)
- **Walking** (FCF10 = 3 or 4, OR FCF11 = 3 or 4, OR FCF14 = 3 or 4, OR FCF15 = 3 or 4) – includes multiple questions on walking with or without equipment
- **Self-care** (FCF16 = 3 or 4)
- **Communication –** must satisfy both conditions below:
  - o Being understood inside household (FCF17 = 3 or 4)
  - o Being understood outside household (FCF18 = 3 or 4)
- **Learning** (FCF19 = 3 or 4)
- **Remembering** (FCF20 = 3 or 4)
- **Concentrating** (FCF21 = 3 or 4)
- **Accepting change** (FCF22 = 3 or 4)
- **Controlling behaviour** (FCF23 = 3 or 4)
- **Making friends** (FCF24 = 3 or 4)
- **Affect –** this domain is measured in two ways: Anxiety and Depression.
  - o **Anxiety** (FCF25 = 1) – measures the frequency of anxiety. Children are considered anxious if they are anxious daily.
  - o **Depression** (FCF26 = 1) – FCF26 measures the frequency of depression. Children are considered depressed if they are depressed daily.

A child aged 5–17 years is considered to have functional difficulties if any of the above variables is present. The following formula is used:

CD=1 if S=1 or H=1 or W=1 or SC=1 or CM=1 or L=1 or R=1 or C=1 or AC=1 or CT=1 or F=1 or X=1 or D=1.

Where:
CD represents children with functioning disability:

- S=1 means the child has functional difficulties associated with seeing
- H=1 means the child has functional difficulties associated with hearing
- W=1 means the child has functional difficulties associated with walking
- SC=1 means the child has functional difficulties associated with self-care
- CM=1 means the child has functional difficulties associated with communicating
- L=1 means the child has functional difficulties associated with learning
- R=1 means the child has functional difficulties associated with remembering
- C=1 means the child has functional difficulties associated with concentrating
- AC=1 means the child has functional difficulties associated with accepting change
- CT=1 means the child has functional difficulties associated with controlling their behaviour
- F=1 means the child has functional difficulties associated with making friends
- X=1 means the child has functional difficulties associated with anxiety
- D=1 means the child has functional difficulties associated with depression

$$SCFD_n = \frac{CD_n}{CD_n + CND_n}$$

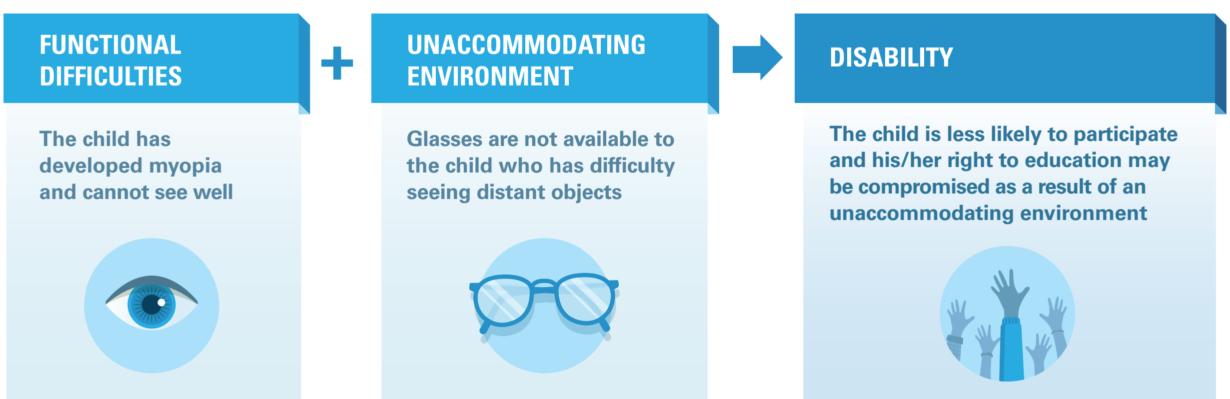
Where:
$SCFD_n$= share of children who have functional difficulties of the n domain
CD_n= children who have functional difficulties of the n domain according to the formula above
CND_n= children who do not have functional difficulties of the n domain according to the formula above

BOX 11: The link between functional difficulties and disabilities

Unaccommodating environments for people with functional difficulties lead to disabilities that often prevent them from making the most of their lives. In the case of education, children with functional difficulties need learning environments that accommodate their needs if they are to succeed in school. An accommodating environment may include an accessible school setting; the provision of necessary equipment such as glasses, hearing aids or learning materials in braille; and well-prepared teaching staff. Such inclusive environments are key to integrating children with functional difficulties in school and ensuring that their right to an education is fulfilled.

A blog post on this topic is available at https://blogs.unicef.org/evidence-for-action/children-disabilities-attend-school-new-findings-sierra-leone.

FIGURE 15 An example showing the link between functional difficulties and disabilities



Source: UNICEF (2019). *Do children with disabilities attend school? New findings from Sierra Leone.* Retrieved from https://blogs.unicef.org/evidence-for-action/children-disabilities-attend-school-new-findings-sierra-leone.

CHAPTER 5

Descriptive analysis

The first step in producing statistical analysis usually consists of descriptive analysis. This type of analysis helps identify the distribution of data and how variables are connected to each other. Descriptive analyses include various summary statistics used to describe statistical data, for example, mean, range, standard deviation, frequency and percentage distribution.

Tabulations and cross-tabulations

A table, or tabulation, is the name given to an arrangement of data in rows and columns. A frequency table, in turn, is a type of table that displays the frequency distribution of the values. Frequency tables are used often because they are easy to interpret. Furthermore, they are important for checking sample sizes in household data.

However, frequency tabulations are unweighted, which makes them unrepresentative of an entire population and less relevant for analysing data sets based on sampling. In the case of household surveys that rely on sampling weights, the most interesting tables for analysis include descriptive statistics other than frequency in the cells.

BOX 12: Sample sizes in MICS reports

In MICS findings reports, some rules are observed regarding sample sizes. If an indicator is calculated based on fewer than 25 unweighted cases, the result is not reported in a table. Furthermore, if the indicator is based on between 25 and 49 unweighted cases, the result is shown in parentheses and should be interpreted with caution.

Source: UNICEF (2019). *Survey Findings Report and Snapshot Guidelines*. United Nations International Children's Fund. New York. Retrieved from http://mics.unicef.org/tools#reporting.

Figure 16 is an example of descriptive statistics showing the share of Iraqi children who are developmentally on track by region (see the ECDI indicator in Chapter 4). These figures are calculated using weights in the questionnaire for children under 5.

FIGURE 16 Share of children 3-4 years old who are developmentally on track in Iraq, by region

Region	Share of children aged 3-4 years old who are developmentally on track
Duhok	87%
Nainawa	71%
Sulaimaniya	94%
Kirkuk	79%
Erbil	88%
Diala	85%
Anbar	82%
Baghdad	84%
Babil	75%
Karbalah	92%
Wasit	71%
Salahaddin	68%
Najaf	76%
Qadisyah	72%
Muthana	91%
Thiqar	63%
Misan	79%
Basrah	75%

Source: Central Statistical Organization/ Kurdish Regional Statistics Office/Ministry of Health/UNICEF (2018). *Iraq Multiple Indicator Cluster Survey 2018, Survey Findings Report*. Baghdad, Iraq: Statistics Iraq.

Tabulations can be decomposed further than two dimensions, if there is interest in seeing how two characteristics interact. These tables are called cross-tabulations. Figure 17 shows the ECDI value by region, disaggregated by age and sex. It can be seen that four-year-old children everywhere in Iraq are more often on track than three-year-olds. The table also shows that in some regions, boys are more on track than girls, while in other regions the situation is reversed.

FIGURE 17 Share of children 3-4 years old who are developmentally on track in Iraq, by age, sex and region

Region	Share of children aged 3-4 years who are developmentally on track			
	Age 3		Age 4	
	Boys	Girls	Boys	Girls
Duhok	85%	78%	89%	95%
Nainawa	69%	74%	65%	78%
Sulaimaniya	92%	90%	100%	95%
Kirkuk	75%	62%	86%	88%
Erbil	82%	86%	95%	88%
Diala	77%	86%	85%	88%
Anbar	78%	88%	70%	94%
Baghdad	80%	85%	87%	85%
Babil	69%	75%	76%	79%
Karbalah	89%	94%	92%	94%
Wasit	68%	75%	75%	68%
Salahaddin	60%	67%	73%	73%
Najaf	75%	85%	72%	75%
Qadisyah	68%	76%	76%	67%
Muthana	89%	90%	95%	89%
Thiqar	66%	59%	73%	46%
Misan	89%	91%	71%	65%
Basrah	62%	82%	78%	80%

Source: Central Statistical Organization/ Kurdish Regional Statistics Office/Ministry of Health/UNICEF (2018) *Iraq Multiple Indicator Cluster Survey 2018, Survey Findings Report*. Baghdad, Iraq: Statistics Iraq.

Correlations and scatter plots

Correlations measure the statistical relationship between two variables in a given data set. Correlations are useful because they indicate how two variables are connected. In some cases, correlations can also show causation, or a predictive relationship. For example, if wealth is correlated with years of education, one could predict that more educated people will be wealthier, or that wealthier people usually receive more education. In this case, as is often the case with correlation analyses, it is impossible to infer the presence of a causal relationship.

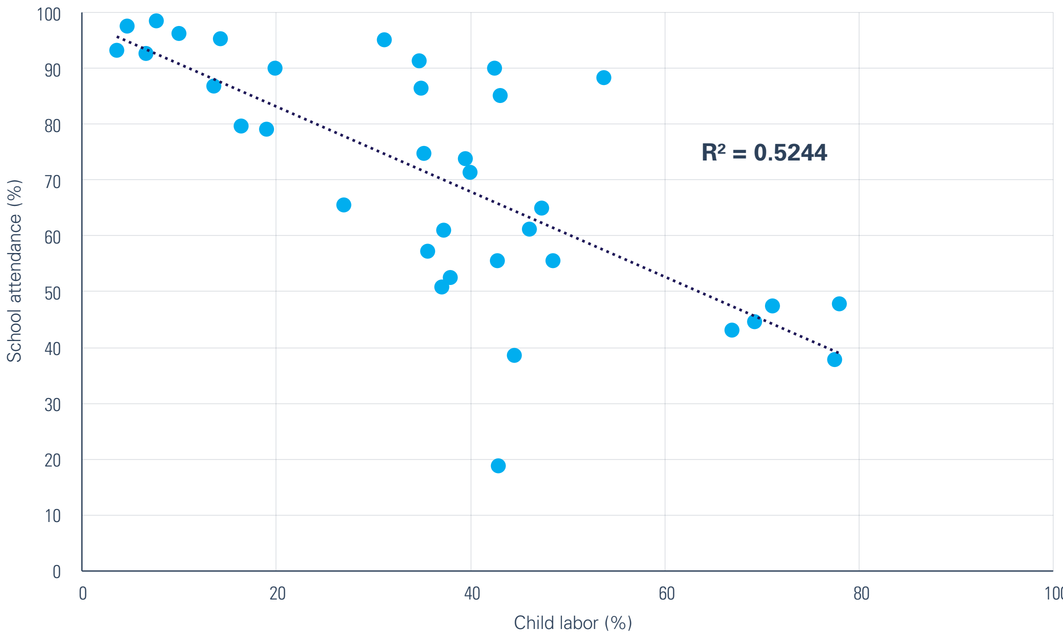
A very useful way of visualizing correlations is through a scatter plot, which is a diagram that typically displays values for two variables from a data set. The data are displayed as a collection of points, with the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis.

Figure 18 uses a scatter plot to show the relationship between child labour and school attendance (1999–2000). The line shows the nature of the relationship is negative: when the rate of child labour increases, school attendance decreases. Therefore, countries with low child labour rates typically have high school attendance rates – as seen in the countries near the top of the figure, including Lebanon and Trinidad and Tobago, whereas countries near the bottom – e.g., Niger, Chad, Guinea-Bissau – display a high prevalence of child labour and low school attendance.

FIGURE 18

Child labour and school attendance for children aged 7–14 years (1999–2000)

Source: Huebler, F. (2008). *Child labour and school attendance: Evidence from MICS and DHS surveys.* Seminar on Child Labour, Education and Youth Employment. Seminar on child labour, education and youth employment. Madrid.



Sub-national disaggregation

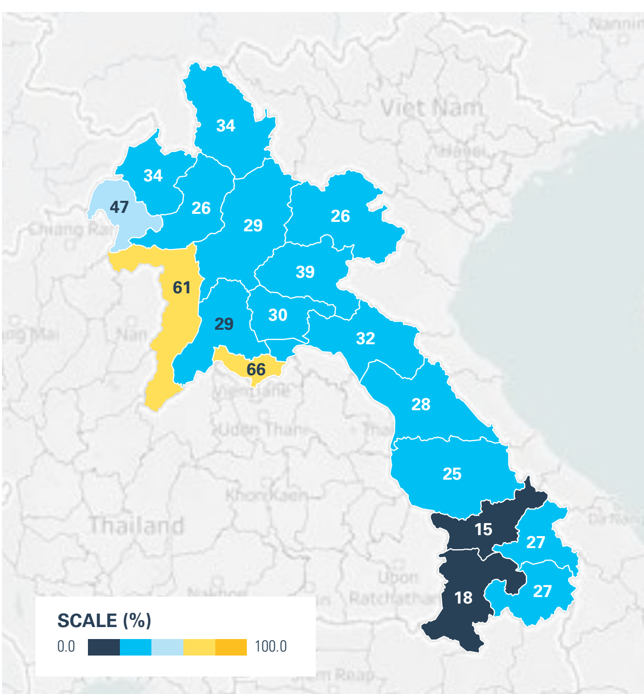
On average, performance across education indicators usually varies by region in a country. Regional analysis is helpful in establishing how regions within a country fare compared to each other. Such an examination helps policymakers serve local needs by identifying region-specific challenges. Tabulations, bar charts and other graphics can display education indicators by region. However, the most visually impactful way to convey regionally disaggregated data is through maps. There are many advantages to such an approach: It can be done for all indicators for which regional data are available, it is visually appealing and it is a quick way to convey a message. Various software can be used to create maps, such as ArcGIS, Infogram, Power BI, Tableau and R (both leaflets and maps). Users must consider the pros and cons of each when choosing the most appropriate software for the type of map they intend to produce.

In Lao PDR, regional disaggregation indicates that ECE attendance varies considerably across provinces as shown in Figure 19. While ECE attendance is strong in the north, education policies for the southern provinces need more attention as the region falls behind in ECE participation.

Fewer than one in five children aged 5 years are attending ECE facilities in Champasak and Saravane Provinces, while two-thirds are attending such facilities in the capital area of Vientiane. Overall, ECE is much more prevalent in the north, where it reaches over one quarter of children in the year before they turn 6, the statutory age for entering primary education.

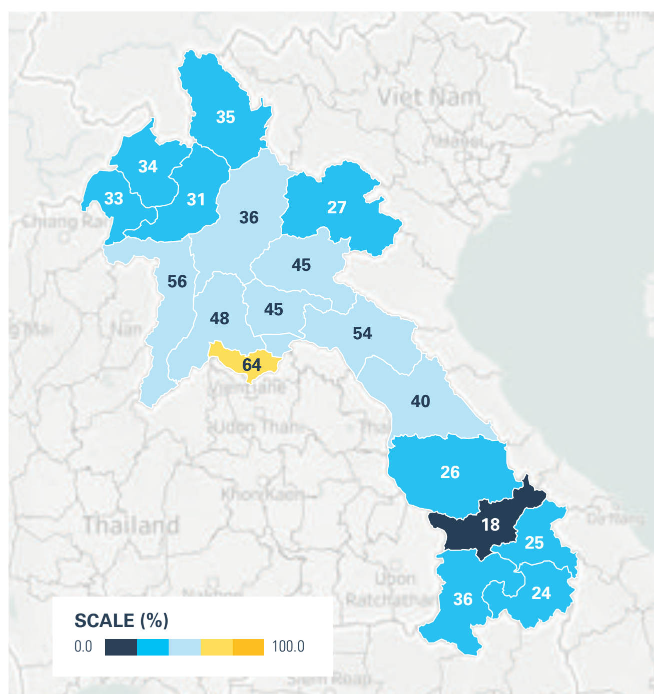
Sub-national disaggregation of ANAR for upper secondary education in Lao PDR, shown in Figure 20, also reveals strong inequalities. While access to upper secondary education is equal to or higher than 40 per cent in all provinces of the country's central region, upper secondary seems to be a bottleneck in the southern and northern regions, which fall behind the centre with upper secondary ANARs below 40 per cent. Saravane Province in the south is an outlier, with an upper secondary ANAR of just 18 per cent, well below that of any other southern province. Overall, access to education is the strongest in the centre and weakest in the south across the two levels analysed.

FIGURE 19 **Early childhood education attendance rate in Lao PDR, by province**



Source: Lao Social Indicator Survey II 2017.

FIGURE 20 **Upper secondary ANAR in Lao PDR, by province**



Source: Lao Social Indicator Survey II 2017.

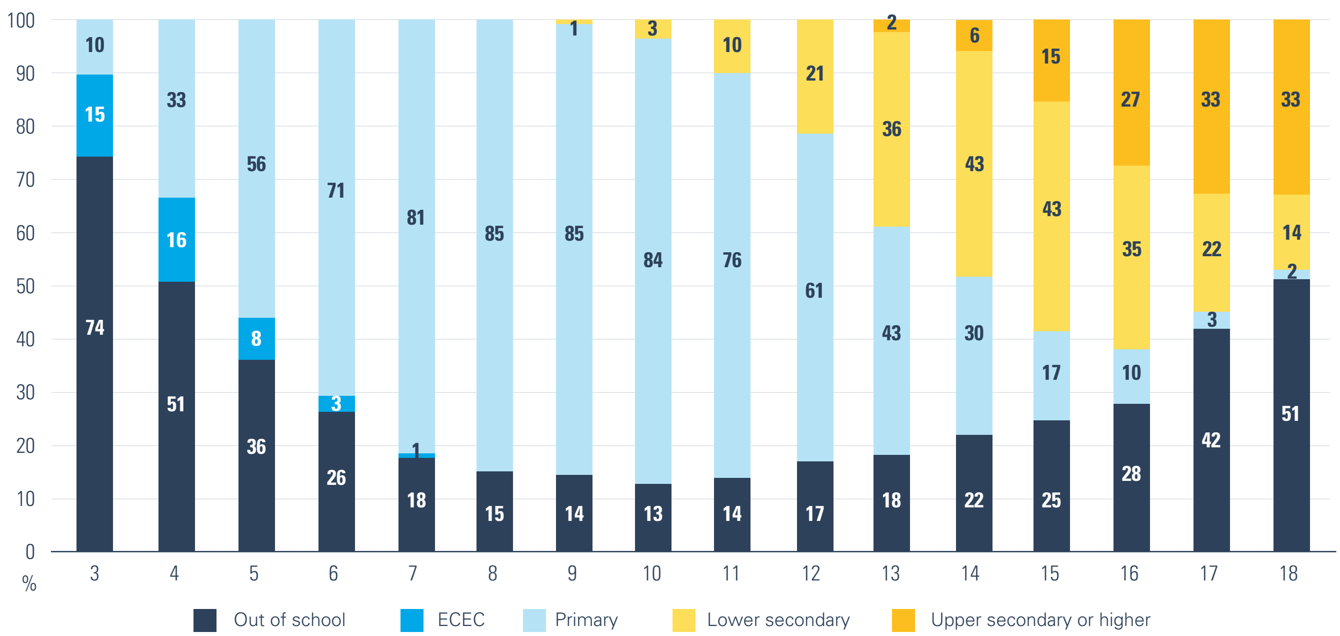
Education status analysis by age cohorts

It is very important to follow children's participation in school as they get older. Education analysis by age cohort helps identify the point at which children enter and leave school. Analysis by cohort is rarely present in statistical reports and cannot be calculated using administrative data. As a result, household data offer valuable information regarding the education status of children of various ages. Figure 21 shows the share of out-of-school children in Sierra Leone for each age and which level of education those in school are attending. Furthermore, the chart helps reveal the stage at which children enter each level of education and when they drop out.

Although the official age to start primary education is 6, the majority of children are already in primary school before they turn 5. In fact, ECE, which was designed for children aged 3 to 5, is rare, peaking at 16 per cent attendance among students aged 4. Primary education, on the other hand, reaches most students aged 5–12 years.

Figure 21 also shows that, although the official age to enter lower secondary education is 12, only 21 per cent of 12-year-olds attend that level of education, while 61 per cent are still in primary school and 17 per cent are out of school. As they grow older, an increasing number of children enrol in lower secondary schools, reaching 43 per cent of those aged 14 or 15. Similarly, entrance into upper secondary education, which should happen when children turn 15, is often delayed and only one-third of children will attend this level of education by the time they are 17 or 18 years old. Finally, the number of children out of school increases for every cohort from age 11, to the point that the majority of 18-year-olds in Sierra Leone are not attending any type of formal education.

FIGURE 21 Education attendance in Sierra Leone, by age



Source: Author's own calculations using Sierra Leone MICS 2017.

Pathway analysis

Pathway analysis is a more extensive version of completion rate analysis. An upper secondary school-age sample is analysed to visually present the students' historical education experience and show their lifetime trajectory in the education system. The reason to choose an upper secondary education aged sample is that these individuals are at the end of their schooling years. The summary chart of a pathway analysis represents how the in-school population gradually shrinks as children progress through the education system, providing new insights to issues of access, completion, repetition, and dropouts. The upper secondary school-age sample is divided into 12 categories, which are presented as pathway bar charts in Figure 22.

Comparing the share of children who entered primary with those who transitioned to upper secondary reflects the efficiency of the school system in ensuring children remain in school and progress in a timely manner.

Figure 22 provides a snapshot of the school trajectories of upper secondary age adolescents in Sierra Leone. For example, the fourth bar shows that children who completed primary school (68 per cent, shown in the third bar) can only fall into three categories: those who completed lower secondary (34 per cent), those who are still attending lower secondary (29 per cent) and those who did not transition to lower secondary (5 per cent).

The pathway analysis of upper secondary school age children in Sierra Leone shows that more than four-fifths of its school-age children are in primary school (82 per cent in the top bar). However, fewer than half of those children transition to upper secondary (32 per cent in the bottom bar). Although some children drop out and some graduate and do not start the next level of education, the biggest group that fails to transition in time are those children still attending lower secondary (29 per cent in the second bar from the bottom) despite being the appropriate age to be in upper secondary school. This points to the need to fight repetition as well as late enrolment to ensure that children enter upper secondary education at the appropriate age.

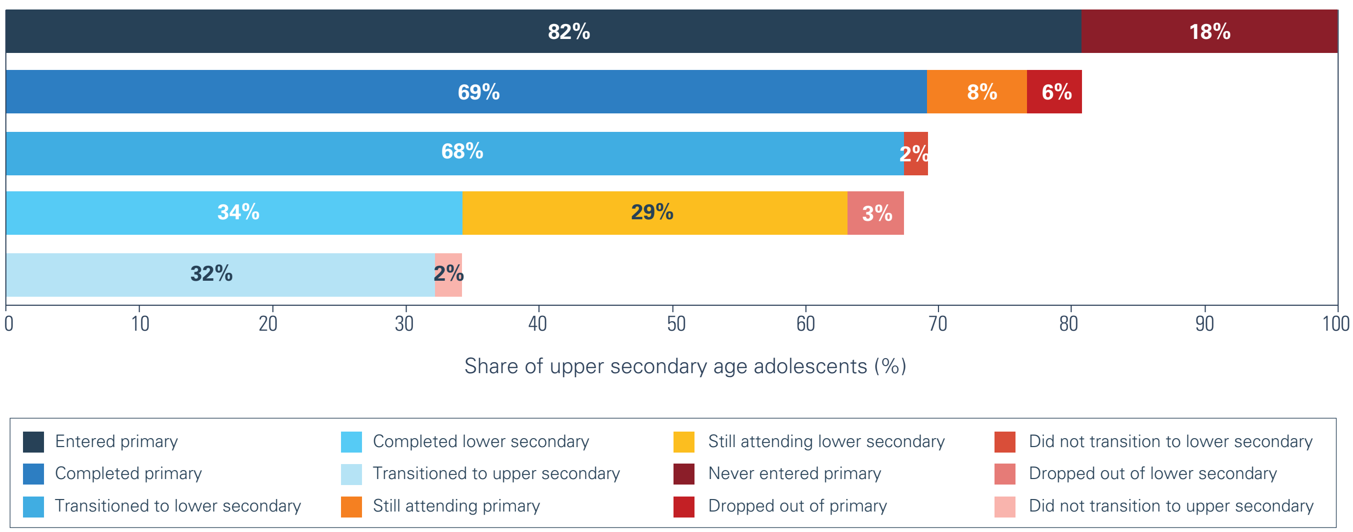
Pathway analysis is also useful for decomposing socioeconomic or ethnic groups within a country. The analysis in Figure 23 shows how the poorest and richest children in Sierra Leone move from the beginning of their education to the transition into upper secondary schools.

When contrasting the poorest and richest quintiles, it is clear that children of upper secondary school age from

the richest households are much more advantaged in each level of education. Wealthier children are more likely enter school, remain in school and graduate from school. However, even for richer children, many students remain behind in lower secondary school when they should already be attending upper secondary. Strikingly, the problem in access and retention happens much earlier among the poorest children in Sierra Leone. About 41 per cent of poor children fail to even enter primary school, and about one in every six who enter primary education drop out. This translates into only 32 per cent of poor children transitioning to lower secondary school and even fewer transitioning into upper secondary education.

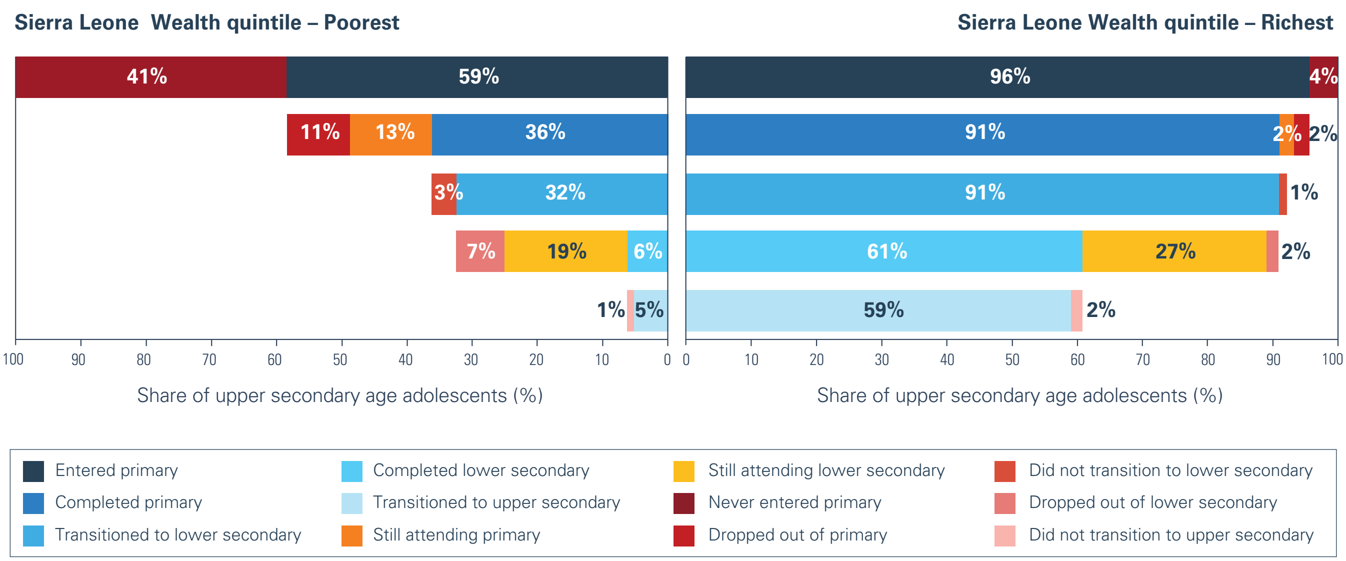
Findings like these are much more easily condensed and conveyed through this type of analysis, which allows for a clear visualization of how different groups in a country progress from primary to upper secondary education.

FIGURE 22 Pathway analysis for Sierra Leone



Source: UNICEF (2018). Education Pathway Analysis dashboard. Retrieved from https://data.unicef.org/resources/education-pathway-analysis-dashboard.

FIGURE 23 Pathway analysis for Sierra Leone, by wealth quintile



Source: UNICEF (2018). Education Pathway Analysis dashboard. Retrieved from https://data.unicef.org/resources/education-pathway-analysis-dashboard.

CHAPTER 6

Regression analysis

Multiple factors can affect education attendance, completion and performance simultaneously, making it challenging to understand which factor in particular is having the stronger effect. For example, if most rural children belong to the poorest segments of wealth distribution and both rural children and children from poorer families are not performing well in school, descriptive statistics are insufficient to point to rurality or poverty as being the most important determinant of poor performance.

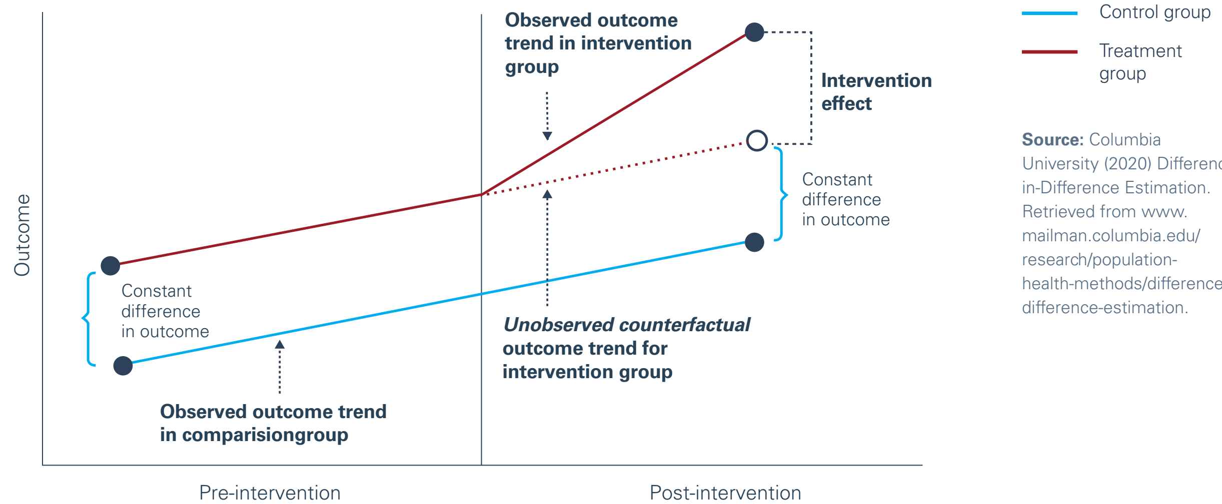
In this situation, regression analysis is used to unpack the influence of each individual factor (e.g., gender, socioeconomic background, ethnicity) on a given phenomenon (e.g., education attendance or performance).

Regressions provide answers to questions about, for instance, the relative importance of a child's wealth in their education attendance when also considering their place of residence as a possible determinant factor. While there are many types of regression analysis, this chapter focuses on logistic models as they are most relevant for analysing education data. It should be noted that specific applications for impact evaluation such as instrumental variable and DID methods will be required to explore causality through regression models. While MICS collects data at a single point in time, other household surveys collect panel data at various points over time. When working with panel data, regression analysis can be performed using the DID technique.

BOX 13: Difference-in-difference

Difference-in-difference (DID or DD) is a statistical technique for observational data that studies the differential effect of an intervention (or treatment) on a "treatment group" that received the intervention, versus a "control group" that did not receive the intervention. It calculates the effect of the intervention (e.g., abolishment of tuition fees in schools, increasing teacher salary, or a change in curriculum) on an outcome (e.g., attendance of a certain group of children or school performance). To assess this effect, the technique compares the average change over time in the outcome variable for the treatment group (shown in red on the chart below), compared to the average change over time for the control group (shown in green below).

FIGURE 24 Illustration of a difference-in-difference estimation



Source: Columbia University (2020) Difference-in-Difference Estimation. Retrieved from www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation.

Logistic models

Variables can be continuous, meaning they can take any value in an interval (e.g., weight, height), or categorical, meaning they describe a defined category and take only certain values (e.g., "attending" or "not attending" school). For example, a student's grades in school are continuous if they are given on a scale between 0 and 100, but they are categorical if students are given letter grades of A, B, C, D, E or F.

Most of the variables used in education data analysis are categorical: e.g., level of education, grade studied, completion and dropout status, attendance, out-of-school status. There are various regression methods to analyse education data, but this guide focuses on the logistic model, which deals with a binary outcome or dependent variable.

A logistic model, also known as logit, aims at estimating the probability of an event. It consists of the estimation of coefficients for each factor explaining a binary phenomenon (e.g., fail or pass, attend or not attend school). Coefficients measure how much a factor explains the phenomenon; positive coefficients mean that the factor has a positive impact on the phenomenon (e.g., the richer one is, the higher the chances of going to school), while negative coefficients show the factor has a negative impact on the measured phenomenon (e.g., the longer hours a child works, the lower the chances of performing well in school). The logistic model will also present statistics to argue whether a coefficient is significantly different from zero or not. Only coefficients significantly different from zero have a considerable positive or negative impact on the phenomenon.

BOX 14: Statistical significance and confidence intervals

Statistical significance is the likelihood that a relationship between two or more variables is caused by something other than chance. Statistical hypothesis testing is used to determine whether the correlation between two variables is statistically significant. This type of statistical test provides a p-value, representing the probability that random chance could explain the result. The closer a p-value is to zero, the more strongly significant a statistical correlation is. In general, most studies aim for p-values under 5 per cent, which mean that the correlation is significant (up to 95 per cent).

The confidence interval of a parameter provides the range of values this parameter can assume within a certain level of significance. For example, a regression model might estimate that a child has a 62 per cent likelihood of being in school based on their individual characteristics. If the confidence interval of this likelihood has a range of 60 per cent to 64 per cent, it means that there is a 95 per cent chance that the probability of a given child attending school will fall between 60 per cent and 64 per cent.

Source: Easton, V. J. and McColl, J. H. (1997). Statistics Glossary v1.1. http://eprints.gla.ac.uk/120164/

Although the direct output estimated by the logistic model is a set of coefficients, these coefficients are hard to interpret for less technical readers. We suggest using two relatively simple outputs of the logistic model: marginal effects and expected values. The marginal effects in logistic regression models present the percentage point difference in the dependent variable (e.g., school attendance, repetition) resulting from changes in independent variables (e.g., gender, rurality, wealth).

For categorical independent variables, marginal effects measure discrete changes in the predicted probability when an independent variable differs between the base category and another category. For example, when explaining school attendance using ethnicity, assuming that ethnicity A is defined as the base, the marginal effect of ethnicity B or C will illustrate how much being of ethnicity B or C changes the probability of attendance in comparison to ethnicity A.

For continuous variables, marginal effects measure the instantaneous rate of change of probability with an extra unit of the independent variable. For example, the marginal effect of wealth measures how much the probability of attending school increases for a child whose family becomes one dollar richer.

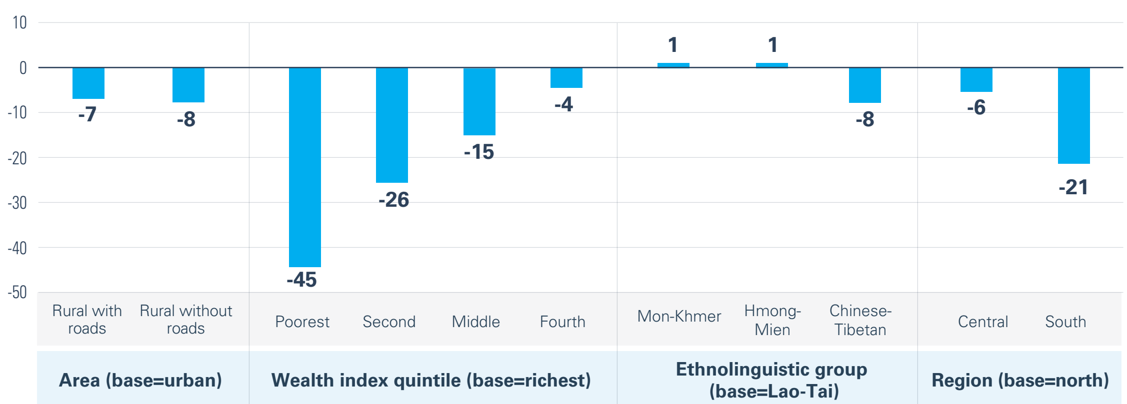
Figure 25 demonstrates the impact each category has in comparison to the base (or reference) category. For example, the base category for location in the figure is urban. This means that, compared to urban children, children from rural areas without roads have an eight-percentage point lower likelihood of attending lower secondary education. The figure shows that socioeconomic status has by far the largest impact in determining a child's lower secondary school attendance after controlling for other factors.

Expected values on the means are calculated to estimate the predicted probability of an independent variable for a specific child, based on the logistic regression performed by inserting mean values for various independent variables. The expected probability of success (e.g., passing a test or attending school) for a child with specific characteristics (e.g., being a girl) will be given by the estimated coefficients, assuming mean values for all the variables other than the category chosen. This means, for example, that the expected value for a girl will be based on the estimated coefficients for girls and the estimated coefficients for other categories, assuming that the girl has the average of each other characteristic (e.g., the mean between rural and urban and across all socioeconomic backgrounds).

For categorical values, the difference between the expected value of one category minus the expected value of the base category equates to the marginal effect (assuming all other variables remain constant on the mean).

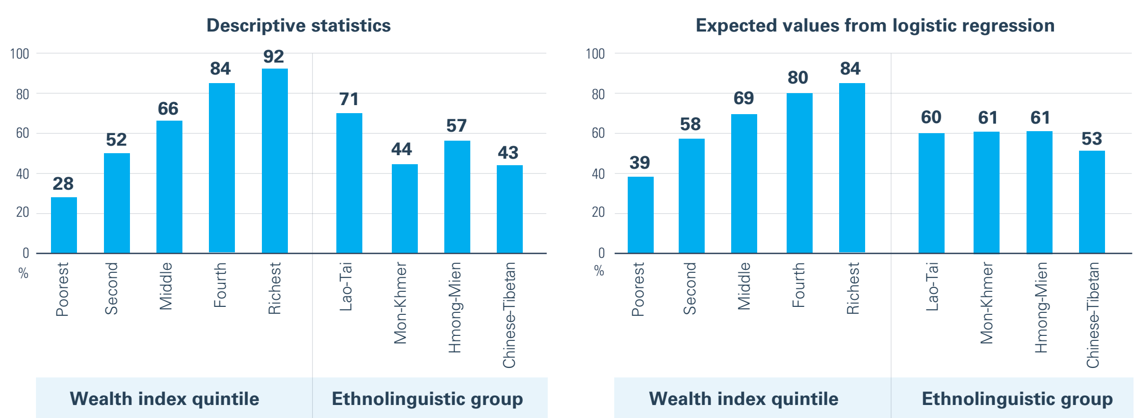
Some groups may appear disproportionately disadvantaged in relation to one another when looking at descriptive statistics, but this can change when examining inferential statistics, such as logistic regressions. The descriptive statistics in the figure below show the ANAR for lower secondary in Lao PDR. It is clear, for example, that the ANAR is much lower for Mon-Khmers (44 per cent) and Chinese-Tibetans (43 per cent) than for the Lao-Tai (71 per cent) ethnolinguistic group.

FIGURE 25 Marginal effects of area, wealth index quintile, ethnolinguistic group and region on lower secondary attendance rates



Source: Author's own calculations using the Lao Social Indicator Survey II 2017.

FIGURE 26 Lower secondary school ANAR in Lao PDR, by sex, area, wealth index quintile, ethnolinguistic group and region



Source: Lao Statistics Bureau (2018). *Lao Social Indicator Survey II 2017, Survey Findings Report.* Vientiane, Lao PDR: Lao Statistics Bureau and UNICEF.

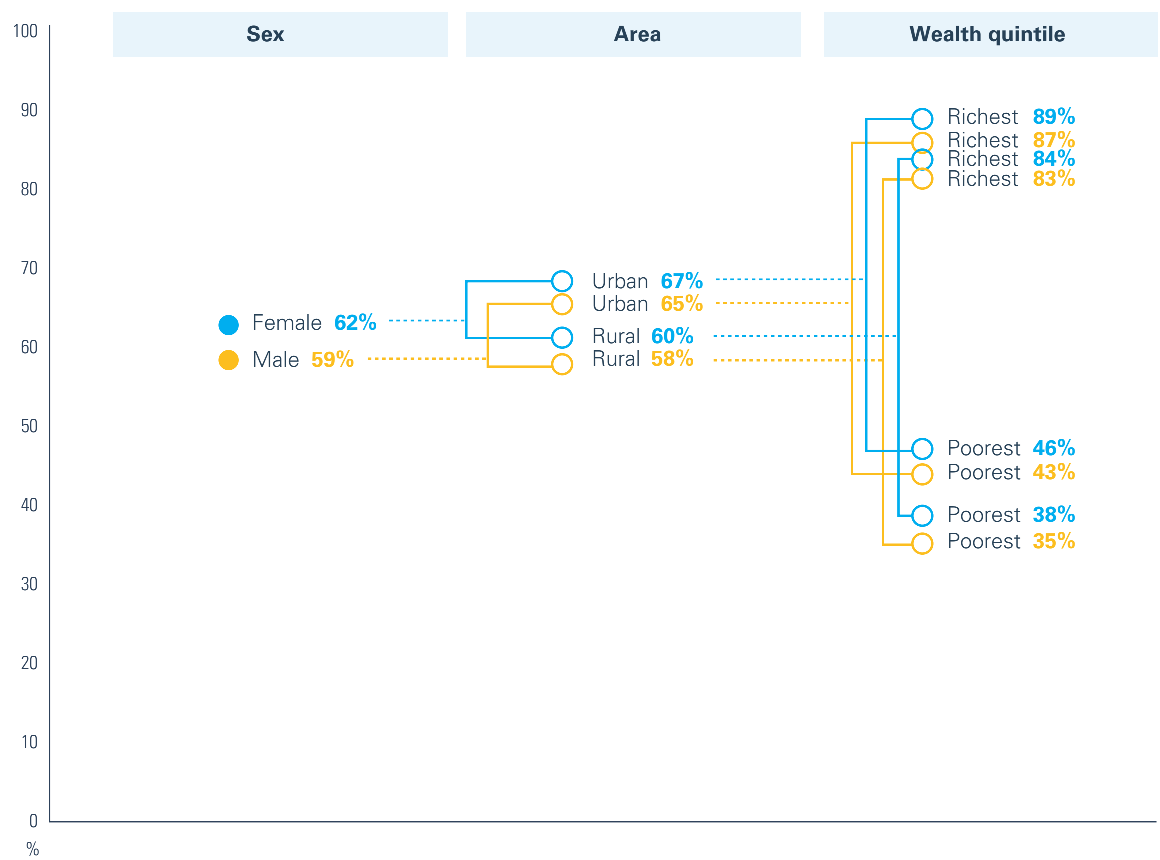
Nevertheless, the level of inequality across ethnolinguistic groups changes when analysing the expected values for attendance, controlling for other factors impacting school attendance at the lower secondary level (such as socioeconomic background).

As shown, the likelihood of attending this level narrows across three ethnolinguistic groups (Lao-Tai, Mon-Khmer and Hmong-Mien). This is due to the fact that much of the difference in attendance across these groups is explained by other factors, namely wealth and region of residency. The regression analysis suggests that what is actually contributing to Mon-Khmer and Hmong-Mien children having worse attendance rates than those from the Lao-Tai group is the fact that they come from poorer families. Even when controlling for sex, area, region and ethnicity, children from the richest families are still twice as likely to be in lower secondary school than those from poorer families.

BOX 15: Probability trees

A graphic way of showing the results of a logistic regression is through a probability tree. The example below shows the probability of a given child attending lower secondary education in Lao PDR according to several of their characteristics. Girls attend lower secondary school in slightly higher numbers than boys, but urban boys are more likely to attend than rural girls. However, when wealth is included as one of the characteristics, it becomes a much more important driver. Boys from rural areas who come from the richest wealth quintile are significantly more likely to attend lower secondary education than girls from urban areas who come from the poorest wealth quintile.

FIGURE 27 Probability tree for lower secondary school attendance in Lao PDR



Source: Author's own calculations using the Lao Social Indicator Survey II 2017.

Multinomial models

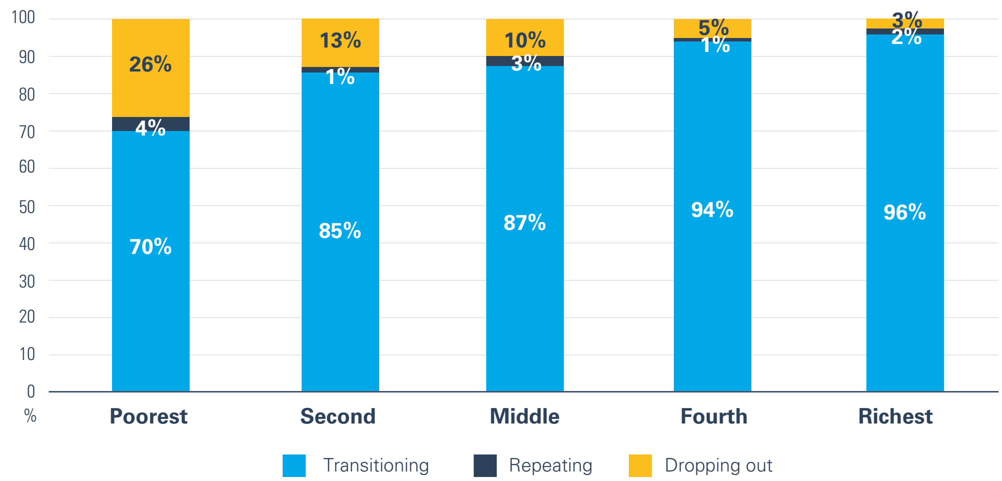
A multinomial model for a logistic regression is a classification method that generalizes logistic regression to more than two binary categories. That is, it is a model used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables. In education data analysis, multinomial models are often used to analyse grade promotion (i.e. the three possible scenarios for a student after a given grade: Promotion to the next grade, repetition, or dropping out) or transition to the next level of the education cycle (e.g., transition from primary to lower secondary education, repetition of the last grade of primary education, or discontinuation of education after graduating primary education).

Figure 28 shows the expected probabilities of students in the last grade of lower secondary education in Lao PDR transitioning to upper secondary, repeating or dropping out, based on the income quintile to which they belong. Students in the poorest quintile have only a 70 per cent chance of transitioning to the next grade compared to the students belonging to the richest quintile, who have a 96 per cent chance. While only a small share of students is expected to repeat a grade across all quintiles, variations between income quintiles reveals that the poorest have the highest probability of both repeating and dropping out.

FIGURE 28

Probability of students in the last grade of lower secondary school in Lao PDR transitioning to upper secondary, repeating or dropping out, by wealth index quintile

Source: Author's own calculations using the Lao Social Indicator Survey II 2017.



Heatmaps

A heatmap is a graphical representation of data where the individual values contained in a matrix are represented as colours. Heatmaps are often used to summarize and present complex results from various regressions. Typically, stronger shades represent larger coefficients, while lighter shades mean coefficients with smaller values or no effect.

Figure 29 presents a heatmap of determinants of school attendance in Lao PDR. It plots three columns with regression results, one for each level of education. For each variable, the category assumed to be the most privileged is omitted from the regression, and the rows for the less privileged categories are populated with dummy (binary) variables. For example, for the five wealth quintiles, there are dummies for the four poorest quintiles, but not for the richest.

The results show blank cells whenever a variable does not significantly impact the results. In Figure 30, girls are not significantly more or less likely than boys to attend primary school, when controlling for all the other variables. When it comes to the gender gap in accessing lower secondary, the results show that girls are 3 percentage points more likely to attend lower secondary schools than boys when other child and household characteristics are held constant. Red cells show characteristics that decrease the odds of children being in school, for example being poor or belonging to the Hmong-Mien ethnolinguistic group. Green cells, in turn, present characteristics that increase the odds of school attendance, such as living in Phongsaly Province.[22]

This sample heatmap illustrates that wealth is the variable that leads to the largest differences in likelihood of attending school. In every level of education, children in the four poorest quintiles (except for the fourth in primary education) are significantly much less likely to attend school than children whose families are in the top wealth quintile. Heatmaps also help to identify other patterns. For example, it is easy to see that wealth is a critical factor associated with access to education across three levels of education, and the magnitude of the effect of wealth on attendance increases for higher levels of education.

FIGURE 29 **Heatmap of marginal effects of various child and household characteristics on school attendance (by percentage point) in Lao PDR**[23]

Variable		Primary	Lower secondary	Upper secondary
Sex	Female		3%	
Area	Rural with roads		-6%	-7%
	Rural without roads		-10%	-15%
Wealth index quintile	Poorest	-10%	-54%	-58%
	Second	-4%	-33%	-47%
	Middle	-2%	-19%	-34%
	Fourth		-6%	-18%
Province	Phongsaly	10%	43%	29%
	Luangnamtha	10%	40%	15%
	Oudomxay	11%	37%	13%
	Bokeo	8%	27%	
	Luangprabang	12%	41%	18%
	Huaphanh	12%	39%	12%
	Xayabury	12%	36%	19%
	Xiengkhuang	12%	47%	23%
	Vientiane	10%	29%	
	Borikhamxay	10%	34%	12%
	Khammua	8%	15%	
	Savannakhet		9%	-9%
	Saravane			-8%
	Sekong	8%	21%	
	Champasack			
	Attapeu	8%	12%	
	Xaysomboune	11%	44%	27%
Ethnolinguistic group	Mon-Khmer	-2%	-5%	-8%
	Hmong-Mien	-7%	-15%	-10%
	Chinese-Tibetan	-3%	-23%	
Mother's education	Post-secondary			
	Upper secondary			-18%
	Lower secondary			-16%
	Primary			-25%
	None or ECE	-5%	-22%	-33%

Source: Author's own calculations using the Lao Social Indicator Survey II 2017.

Cross-sectoral analysis

Cross-sectoral analyses are used to link educational outcomes with characteristics from other sectors. This type of work deepens understanding of, for example, how education attainment and learning are associated with other factors, including child labour, child marriage, home environment and parental involvement. Both descriptive and regression analysis can be useful tools for cross-sectoral work.

Example 1: Child labour

The descriptive cross-sectoral analysis presented in Figure 30 shows that more children who are out of school (47 per cent) engage in child labour in Sierra Leone than those attending school (37 per cent).

Jointly analysing child labour and education can be tricky as some factors can influence both school attendance and child labour. For example, many out-of-school children are more rural and poorer than their counterparts, and children engaged in child labour are also more rural and poorer. As a result, it is important to understand whether children in child labour have lower attendance rates due to child labour itself or to factors affecting both child labour and attendance simultaneously (e.g., poverty and rurality).

Regression analysis can determine if there is a link between school attendance and child labour by helping to identify which factor has the largest impact on school attendance, when other important characteristics are controlled for. It makes it possible to disentangle whether lower attendance is directly linked to child labour or to poorer socioeconomic background – the latter increases chances of both being out of school and working at an early age.

Figure 31 shows the results of regressions for expected attendance for children in Sierra Leone, controlling for several socioeconomic variables such as sex, region of residency and wealth quintile. The differences in likelihood of attending school between working and non-working children increase strongly in significance after age 11. At age 11, the likelihood of non-working child attending school is 85 per cent versus 83 per cent for a working child. At 17, the gap increases substantially – 69 per cent of non-working children attend school, while only 54 per cent of working children do.

FIGURE 30

Children in child labour in Sierra Leone, by area, region, age, school attendance, mother's education and wealth index quintile

Source: Statistics Sierra Leone (2018). *Sierra Leone Multiple Indicator Cluster Survey 2017, Survey Findings Report*. Freetown, Sierra Leone: Statistics Sierra Leone.

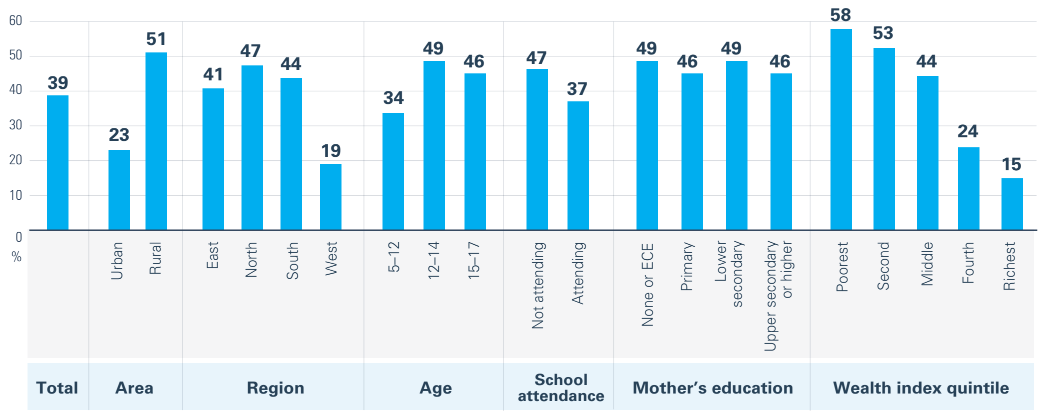


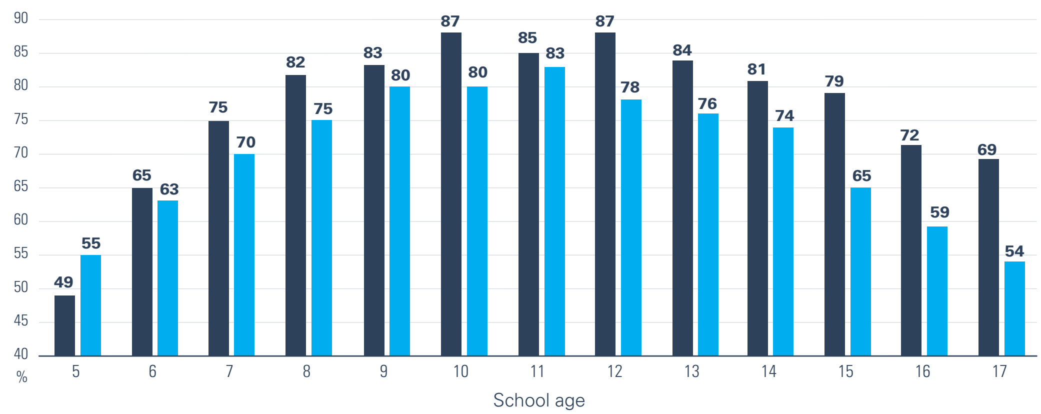
FIGURE 31

Expected attendance of children in Sierra Leone, by age and working status

Not working

Child labour

Source: Author's own calculations using Sierra Leone MICS 2017.



Example 2: Child marriage

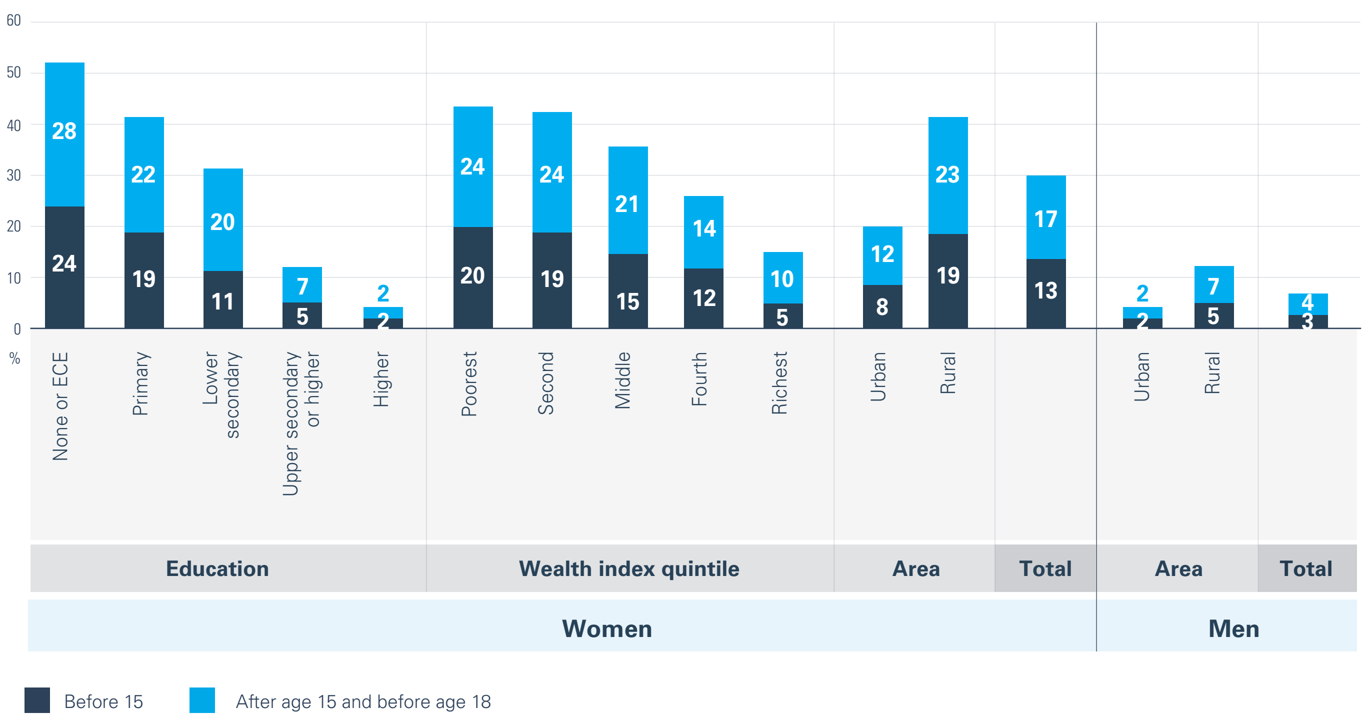
Another area that yields interesting cross-sectoral analytical work is child marriage, which is common in many countries, particularly among girls. Marriage at a young age can hamper progress in school and increase dropout rates. As a result, whenever possible, early marriage should be eliminated, and children should be encouraged to stay in school for as long as possible. Figure 32 uses data from Sierra Leone to investigate the relationship between child marriage and education attainment.

Far fewer boys than girls get married before they turn 15 – only 3 per cent of young males married before the age of 15 and 7 per cent married before the age of 18. For females of the same age, 13 per cent married before they were 15 and a notable 31 per cent married before they were 18. The same factors impacting attendance are also present as determinants of early marriage. Among rural girls, those who are less educated and poorer are more likely to marry early. However, education seems to be the factor responsible for the largest gap between those who marry and those who do not.

The regression analysis in Figure 33 shows how likely a girl is to get married when controlling for various factors at the same time. As mentioned, the use of regression helps determine the correlation between two variables, while also controlling for several individual characteristics. The results show that education is more correlated with early marriage than socioeconomic background. When controlling for sex, area of residence, region, socioeconomic background and highest level of education attended, there is a larger difference between the likelihood of educated and uneducated girls marrying early than in the likelihood of poorer and richer girls marrying early.

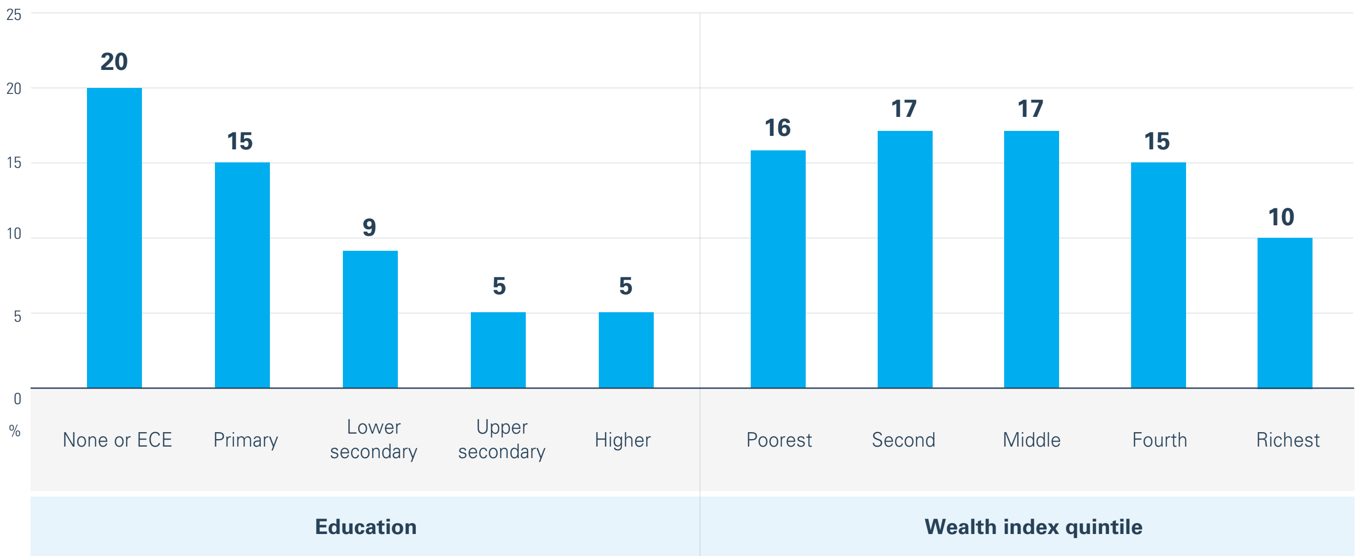
Indeed, only girls in the top wealth quintile are significantly less likely to marry early in comparison to their peers. Nonetheless, lower levels of education are correlated with a higher likelihood of girls getting married early, even when controlling for other socioeconomic factors. Regression analysis shows correlations between variables, and in the case of education, the relationship is reciprocal – education influences early marriage, as much as early marriage influences education. As a result, it can be the case that girls who marry early are more likely to abandon school and less likely to proceed to higher levels of education, but also that girls who stay in school longer are less likely to marry before they turn 15.

FIGURE 32 **Percentage of young people aged 19 to 25 years married before the age of 18 in Sierra Leone, by education, wealth index quintile and area**



Source: Statistics Sierra Leone (2018). *Sierra Leone Multiple Indicator Cluster Survey 2017, Survey Findings Report*. Freetown, Sierra Leone: Statistics Sierra Leone.

FIGURE 33 Percentage of girls who marry before the age of 15 in Sierra Leone, by education and wealth index quintile



Source: Author's own calculations using Sierra Leone MICS 2017.

Example 3: Children with disabilities

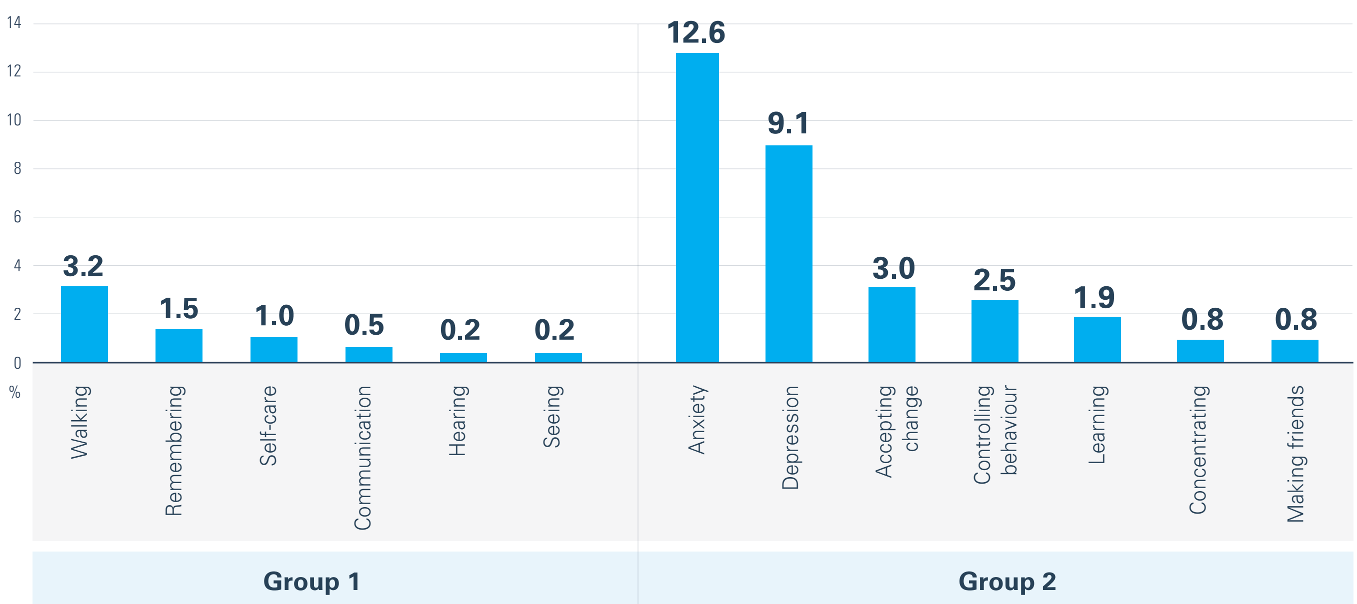
The Washington Group/UNICEF modules on Child Functioning, finalized in 2016, cover children aged 2–4 years and 5–17 years. These two modules assess functional difficulties in different domains, including hearing, vision, communication/comprehension, learning, mobility and emotions. The purpose of these new modules, which are included in almost all the MICS6 household surveys, is to identify the subpopulation of children at greater risk of experiencing difficulties in accessing and participating in school due to systematic issues such as an unaccommodating school environment, social prejudice and so on.

The prevalence of different types of child functional difficulties varies widely. In Sierra Leone, anxiety and depression are the most common, while hearing and seeing difficulties affect only 0.2 per cent of children. Given their various prevalence rates and levels of complexity, disabilities can be divided in two categories for the analysis. The first is associated with the functional domains defined for adults (group 1 in Figure 34), including six functional domains put together in 2001 by the Washington Group on Disability Statistics (seeing, hearing, walking, cognition, self-care and communication).[24] The second group (group 2 in the figure) includes the additional child functional domains of cognitive and behavioural characteristics,[25] measurement of which are important to understanding children's difficulties and promoting inclusive education.

Beyond understanding the prevalence of various functional difficulties, it is also important to understand which functional difficulties schools are more or less prepared to accommodate. Figure 35 shows the likelihood of school attendance using a regression analysis that controls for several variables, including gender, wealth index quintile and area of residence. The higher and lower bounds show the maximum and minimum expected values with a 95 per cent confidence interval, meaning that there is a 95 per cent chance of actual attendance falling between the lower and higher bounds. If the values of a given category fall between the lower and higher bounds of another category, it means that it cannot be assumed that, within a 95 per cent confidence interval, the values for the two categories are significantly different from each other.

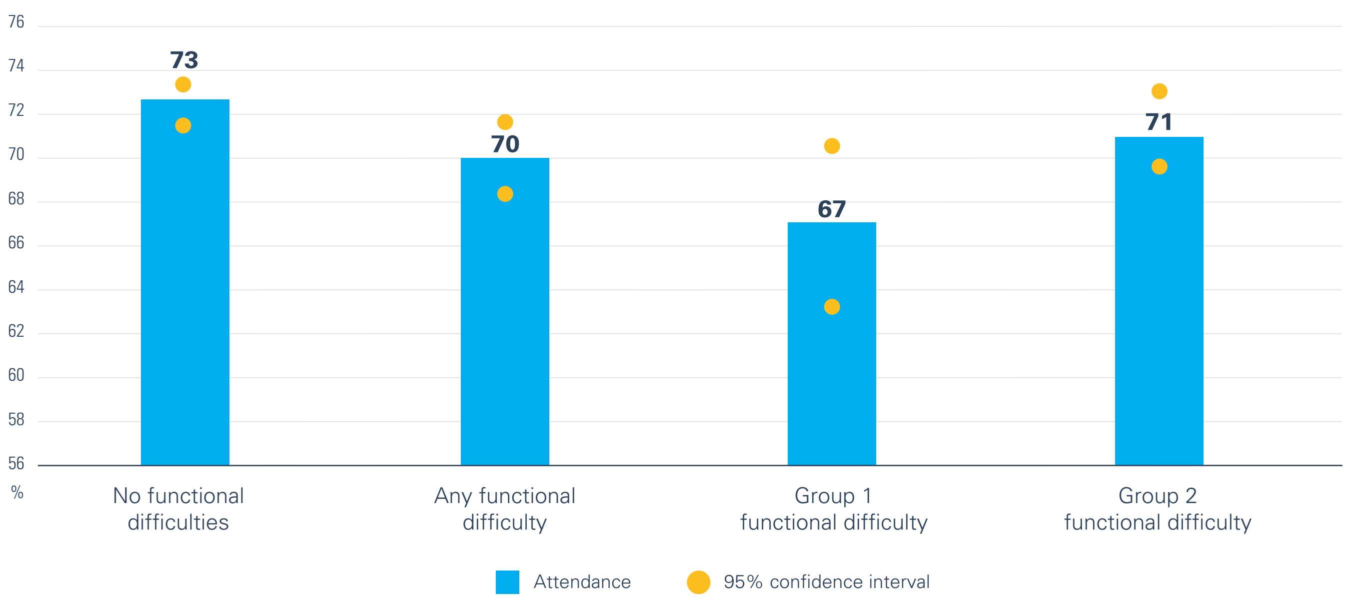
From Figure 35, it can be concluded that children with disabilities are not significantly less likely to attend school than children without disabilities. Nevertheless, schools seem much less accommodating to children with functional difficulties in group 1, who have only a 67 per cent chance of being in school between ages 5 and 17, while children without any functional difficulties have a 73 per cent likelihood, which is significantly different.

FIGURE 34 Percentage of children aged 5–17 years in Sierra Leone with functional difficulty across domains



Source: Statistics Sierra Leone (2018). *Sierra Leone Multiple Indicator Cluster Survey 2017, Survey Findings Report*. Freetown, Sierra Leone: Statistics Sierra Leone.

FIGURE 35 **Likelihood of children aged 5–17 years in Sierra Leone attending school, by functional difficulty status controlling for socioeconomic and demographic characteristics**



Source: Author's own calculations using Sierra Leone MICS 2017.

CHAPTER 7

Case studies

This section presents three case studies that show how indicators can be linked to policy actions.

Case study 1: Promoting early education

Early learning provides tremendous lifelong advantages to children both in terms of their future education and life skills. Early enrolment in age-appropriate programmes allows children to learn at the right age and to be prepared to move forward to higher levels of education.

In Lao PDR, the starting age for primary school is 6 years old. Figure 36 shows that 74 per cent of children in the country who are aged 5 years attend some form of organized learning. Nonetheless, a more careful assessment of the data shows that only 34 per cent of children aged 5 years attend ECE, which is the education level specifically designed for their age group.

It is important that children receive education that is appropriate to their age and that they attend schools prepared to teach age-relevant content, which is why in 2015, ECE was made free of charge in Lao PDR. Despite this measure, many parents continue to enrol their children aged 5 years and younger in primary education, instead of ECE, which is designed for their age group.

Several factors hamper age-appropriate ECE enrolment, including: 1) a lack of access, as ECE programmes may be unavailable in some areas of the country; 2) a misperception among parents about the advantages of early enrolment in primary education instead of ECE; and 3) a heavy focus from parents on academic, instead of holistic, development for their young children.

ECE's impact on learning outcomes

Early childhood education plays a pivotal role in ensuring that children learn and that they transition from the pre-education phase, which fosters holistic child development and active learning, into the education phase, which equips children with the capacity to learn and thrive. Children who are too young to attend primary school should be given the chance to participate in education through ECE, which aids their transition and guarantees that they will be developmentally on track and ready to start primary education at the appropriate age.

As discussed, child development can be measured using a series of variables including social, numeracy and reading skills. One indicator of child development is literacy-numeracy, which consists of a combination of three capacities expected from children: reading four simple words, identifying at least 10 letters of the alphabet and knowing their name, and recognizing numbers from 1 to 10.

FIGURE 36

Education level attended among children aged 3–6 years in Lao PDR

Source: Author's own calculations using the Lao Social Indicator Survey II 2017.

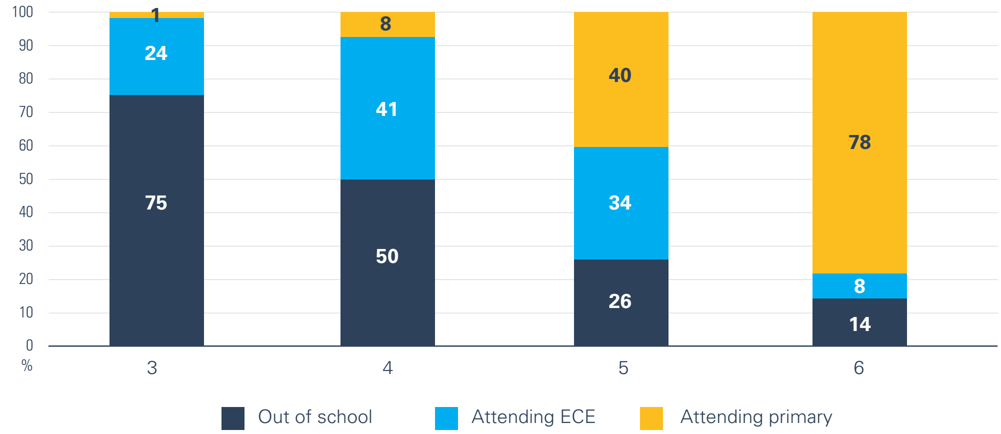
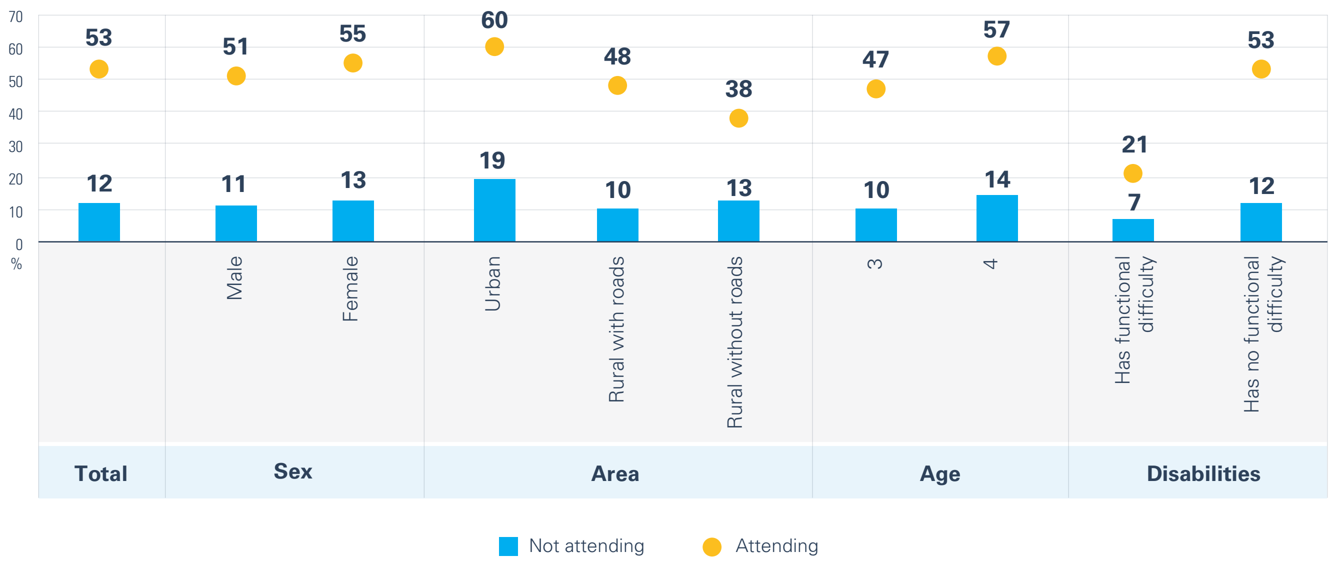


FIGURE 37 Percentage of children aged 3–4 years in Lao PDR developmentally on track in literacy-numeracy, by ECE attendance



Source: Author's own calculations using the Lao Social Indicator Survey II 2017.

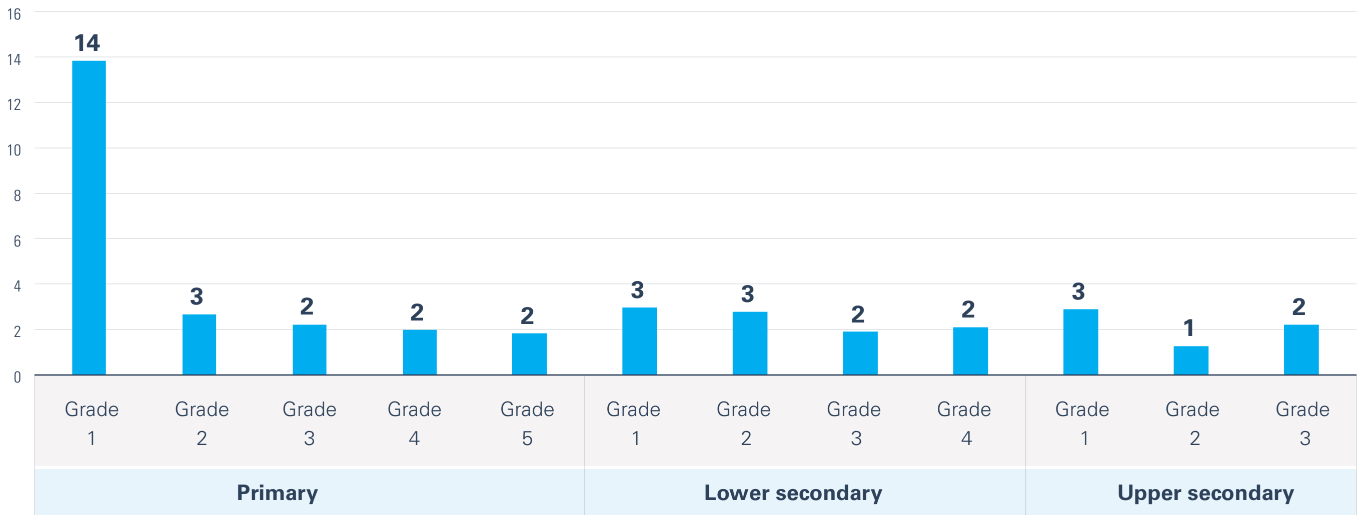
As seen in Figure 37, data from Lao PDR reinforce the need for early enrolment in ECE and the role it has in equipping children with the necessary literacy-numeracy competencies to start primary school. As shown by the "Total" category at far left, nationally, only 12 per cent of children who miss out on ECE meet the level of development of literacy-numeracy skills expected for their age group, which makes them less prepared to start primary education. Meanwhile, children from all backgrounds who attend ECE are much more likely to be developmentally on track (53 per cent).

Repetition and premature enrolment

In Lao PDR, 40 per cent of children in Grade 1 are under-age for primary school. Early enrolment in primary education means that children start Grade 1 when they are younger than the official age of school entry of 6, which is the age schools are prepared to accommodate. Schools, curricula and teaching materials are all designed for students who are of a given age and should be attending the appropriate grade for that age. Students who are too young to attend a certain grade can have more difficulties integrating in school.

Figure 38 shows that repetition in most grades in Lao PDR is low, with the notable exception of Grade 1 of primary education, which a dramatic 14 per cent of students repeat. Starting the first grade of primary school can be a difficult change for many students, which explains higher repetitions at this stage. However, the first grade of primary school is particularly complicated for those who enter primary education too young. Figure 40 shows the distribution of repeaters in Grade 1 by age to explain the importance of age-appropriate learning. Many more children who started too young at age 4 or 5 repeat Grade 1 than do those who started at age 6, 7 or 8. This means that younger children have a harder time adapting to primary school. One contributing factor could be that those children who entered primary education too young did not previously attend ECE and were not equipped with the literacy-numeracy skills that ECE provides.

FIGURE 38 Grade repetition per grade in Lao PDR, by level of education



Source: Author's own calculations using the Lao Social Indicator Survey II 2017.

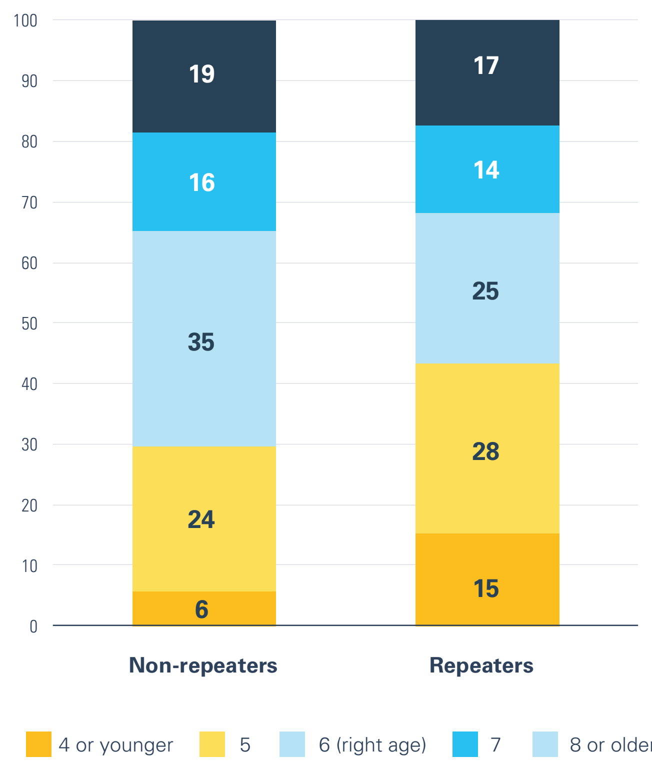
Measuring early learning and expanding ECE

Improving access to ECE to equip children with the right competencies before they start primary education, as well as to prevent repetition early on in primary education, should be a priority. Sometimes a lack of ECE availability pushes parents to send children to primary school when they are too young to attend. However, early primary school attendance hampers learning and leads to higher repetition among younger children in early grades.

Besides repetition, there is also a strong negative link between sending children to primary school too early and the quality of primary education. Children aged 5 years are too young to attend primary school, but Figure 36 shows that in Lao PDR, a total of 40 per cent do anyway. This influx of under-age children increases the number of students per teacher in Grade 1 and lowers the quality of education at this critical transition from ECE to primary education. Having more children in class, especially given that some of the students are not mature enough to pass that class, decreases the quality of education for all students and increases the likelihood of repetition.

ECE has a strong impact on ensuring that children are developmentally on track for their age and that they are prepared to start primary education with the necessary skills. It is important that schools providing ECE are available and that all children succeed in attending and getting ready to start primary education at the appropriate age.

FIGURE 39 Repeaters of Grade 1 in Lao PDR, by age



Source: Author's own calculations using the Lao Social Indicator Survey II 2017.

Case study 2: Do children with disabilities attend school?

As access to education increases around the world, so do concerns about schools providing the necessary support to children with disabilities. Unaccommodating environments for students with functional difficulties can often prevent them from making the most of their education.

For children with functional difficulties to succeed in school, learning environments should accommodate their needs. An accommodating environment may include an accessible school setting, the provision of necessary equipment such as glasses, hearing aids or learning materials in braille, and well-prepared teaching staff.

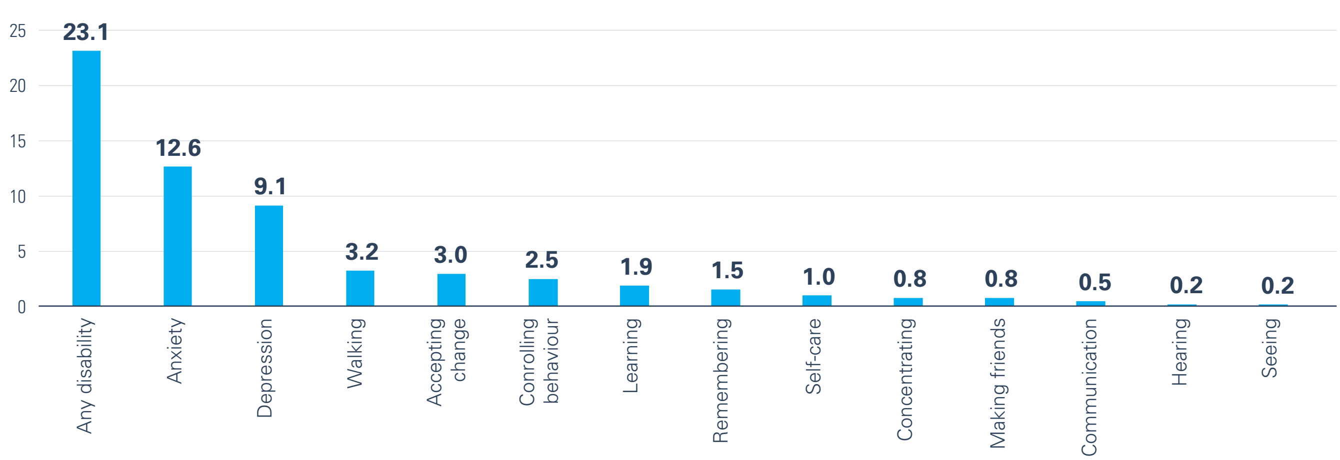
Unaccommodating environments prevent these children from being able to enjoy the same opportunities as others. Unaccommodating environments lessen the chances that children with functional difficulties will get an education and inhibits their participation in the learning experience. In such environments, parents may be reluctant to send children with functional difficulties to school due to concerns that the appropriate resources are not in place to support learning. Some parents may also make this decision out of fear that their children will be bullied or socially excluded.

In many countries, data on these disadvantaged children and how they participate in school often lack quality and comparability. Sierra Leone's MICS revealed that about 23 per cent of children aged 5–17 years have at least one functional difficulty, as seen in Figure 40. However, prevalence varies widely across functional domains. Based on information from their mothers or primary caregivers, around 13 per cent of children display signs of severe anxiety and around 9 per cent are very sad or depressed on a daily basis. In contrast, only 0.2 per cent of children in Sierra Leone have difficulties seeing or hearing due to the lack of appropriate equipment or because glasses and hearing aids are not effectively addressing their needs.

The findings further demonstrate that many children with functional difficulties are attending school in Sierra Leone. For example, Figure 41 shows that children with functional difficulties in managing emotions (depression and anxiety) attend schools at similar rates to those without functional difficulties. In contrast, the conventional view that children with disabilities are more likely to be out of school still holds for other functional domains. For example, the survey found that children with seeing or hearing difficulties had the lowest school attendance rates of 60 and 50 per cent, respectively.

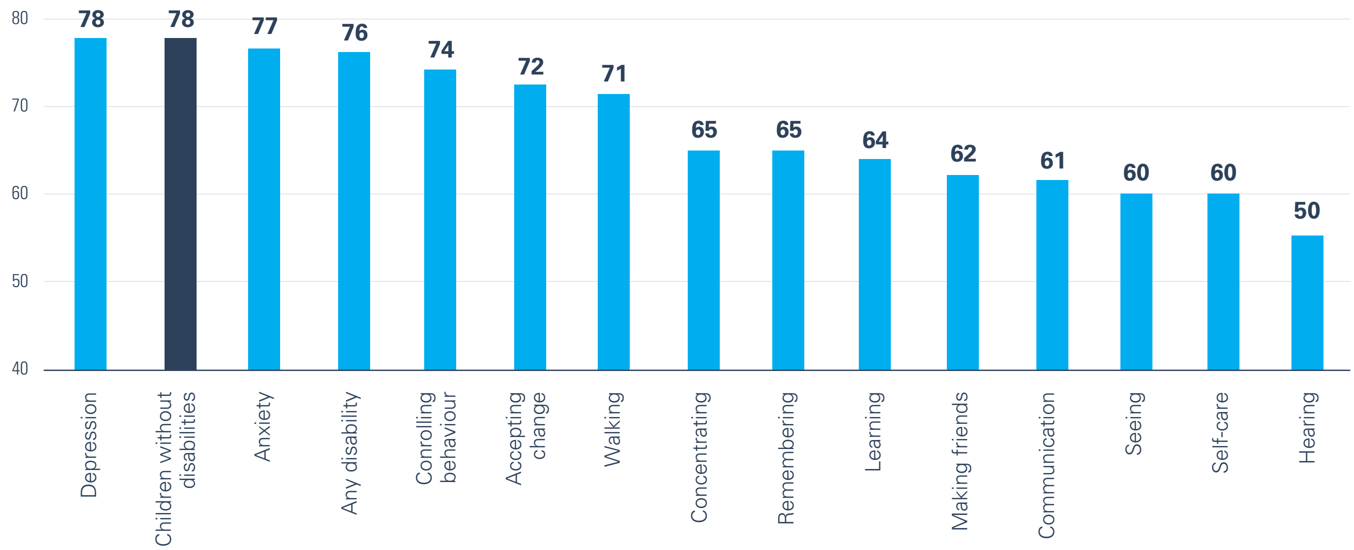
The results indicate that some traditional ideas about children with disabilities do not always prove to be true. On one hand, most children with disabilities are already in school, but ways to meet their unique needs and efficiently educate them need to be put in place or to be improved. On the other hand, children with difficulties such as seeing or hearing continue to be left out, and have problems accessing education. This indicates a need to further invest in accommodating environments to fulfill the needs of all children in the classroom, as well as to develop more inclusive education policies to help all children enter and remain in the school environment.

FIGURE 40 **Prevalence of disability among children aged 5–17 years in Sierra Leone, by functional domain**



Source: Statistics Sierra Leone (2018). *Sierra Leone Multiple Indicator Cluster Survey 2017, Survey Findings Report.* Freetown, Sierra Leone: Statistics Sierra Leone.

FIGURE 41 Percentage of children aged 5–17 years in Sierra Leone attending school, by disability status



Source: Author's own calculations using Sierra Leone MICS 2017.

Case study 3: How much do children learn?

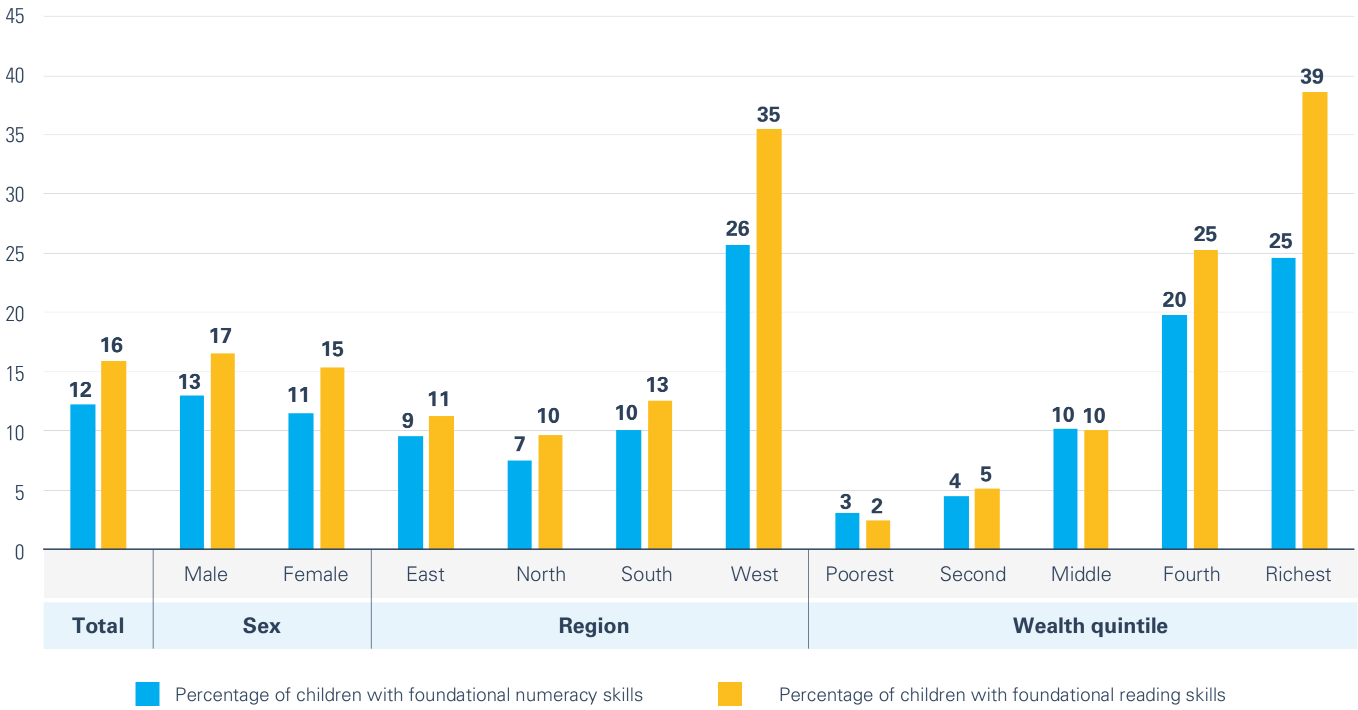
If a child attends school every day, can we assume that he or she is learning? An increasing number of children around the world have gained access to education, particularly in low-income countries, but until recently, it has been unknown whether more children are learning as a result.

Findings show that a mere 12 per cent of children in Grades 2 and 3 in Sierra Leone meet the expected levels of numeracy skills for their grade, but not all children fare the same. How a child performs is strongly defined by where they grow up and by the wealth of their parents, as there are strong regional and socioeconomic inequalities in the country. As shown in Figure 42, in the western part of Sierra Leone, where the capital is located, three times more children achieve the expected reading skills for their grade than in the rest of the country. Among the richest children, around 39 per cent demonstrate basic literacy skills, while only an alarming 2 per cent of poor children do.

Virtually all children out of school fail to display foundational skills. Among those in school, only 8 to 9 per cent of children in Grade 3 have the basic reading and math skills expected for that education level. Despite showing low levels of learning at the beginning of primary education, as children in Sierra Leone move on to higher grades their learning performance improves, although it remains lower than expected. By lower secondary level, more children display the expected foundational skills for Grades 2 and 3, although many do not. In the first grade of lower secondary education, only 66 per cent of children have the reading skills they should have achieved by Grades 2 and 3 and just 42 per cent have the mathematical skills expected for these grades. This speaks to the importance of investing in quality primary education, particularly in earlier years, to ensure that students are equipped with foundational skills at the right age.

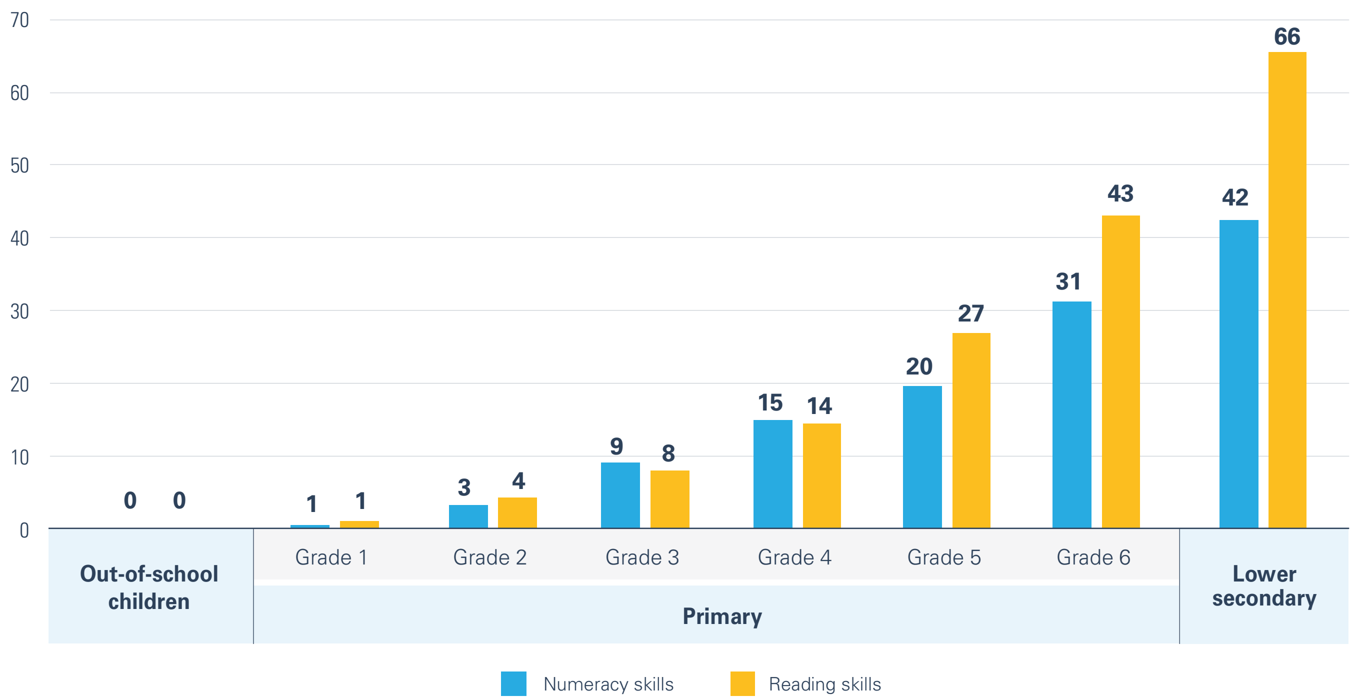
Though many children in Sierra Leone are learning, they appear to be missing out on a crucial window of opportunity during the early years of life, potentially denying them cascading advantages as they grow into adulthood. As illustrated in Figure 43, new MICS data show that many children fail to learn even when they stay in school for several more years. Although substantial efforts have been made to guarantee that every child everywhere goes to school, it is essential that we start focusing on learning outcomes as well. Access to education is key to ensure learning, given that children out of school systematically fail to acquire foundational skills. However, the quality of schools is also extremely important, particularly in ECE and primary school where children are equipped with the knowledge and skills needed to thrive.

FIGURE 42 **Percentage of children in Grades 2/3 who demonstrate foundational skills in Sierra Leone, by sex, region and wealth index quintile**



Source: Statistics Sierra Leone (2018). *Sierra Leone Multiple Indicator Cluster Survey 2017, Survey Findings Report*. Freetown, Sierra Leone: Statistics Sierra Leone.

FIGURE 43 **Percentage of children in Sierra Leone who demonstrate foundational skills in numeracy and reading, by grade attended**



Source: Author's own calculations using Sierra Leone MICS 2017.

BIBLIOGRAPHY

Central Statistical Organization/Kurdish Regional Statistics Office/Ministry of Health/UNICEF (2018). *Iraq Multiple Indicator Cluster Survey 2018, Survey Findings Report*. Baghdad, Iraq: Statistics Iraq.

Columbia University (2020). "Difference-in-Difference Estimation." Retrieved from www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation

Dayıoğlu M. (2012). *How sensitive are estimates of working children and child labour to definitions? A comparative analysis*. MICS Methodological Papers, No. 1, Statistics and Monitoring Section, Division of Policy and Strategy. New York: UNICEF.

Easton, V. J. and McColl, J. H. (1997). *Statistics Glossary v1.1*. Retrieved from http://eprints.gla.ac.uk/120164/

Gochyyev P., Mizunoya S., and Cardoso M. (2019). *Validity and reliability of the MICS foundational learning module*. MICS Methodological Papers, No. 9, Data and Analytics Section, Division of Data, Research and Policy. New York: UNICEF.

Huebler, F. (2008). *Child labour and school attendance: Evidence from MICS and DHS surveys*. In Seminar on Child Labour, Education and Youth Employment.

Hattori H., Cardoso M., and Ledoux B. (2017). *Collecting data on foundational learning skills and parental involvement in education*. MICS Methodological Papers, No. 5, Data and Analytics Section, Division of Data, Research and Policy. New York: UNICEF.

Lao Statistics Bureau (2018). *Lao Social Indicator Survey II 2017, Survey Findings Report*. Vientiane, Lao PDR: Lao Statistics Bureau and UNICEF.

Loizillon, A., Petrowski N., Britto P., and Cappa C. (2017). *Development of the Early Childhood Development Index in MICS surveys*. MICS Methodological Papers, No. 6, Data and Analytics Section, Division of Data, Research and Policy. New York: UNICEF.

Project ROSALIND (2020). Glossary: Probability Tree Diagram. Retrieved from http://rosalind.info/glossary/probability-tree-diagram

Ranguelov S. (2010). *International Standard Classification of Education (ISCED) 2011*. Presented at the Eurydice working party meeting.

Schroeder, L. D., Sjoquist, D. L., and Stephan, P. E. (2016). *Understanding regression analysis: An introductory guide* (Vol. 57). Sage Publications.

Starkweather, J., and Moske, A. K. (2011). *Multinomial logistic regression.*

Statistics Sierra Leone (2018). *Sierra Leone Multiple Indicator Cluster Survey 2017, Survey Findings Report*. Freetown, Sierra Leone: Statistics Sierra Leone.

Triola, M. F., Goodman, W. M., Law, R., and Labute, G. (2006). *Elementary statistics*. Reading, MA: Pearson/Addison-Wesley.

UNESCO-UIS (2012). *International Standard Classification of Education: ISCED 2011.* Paris: UNESCO. Retrieved from http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

UNESCO-UIS (2017). *Fact sheet number 47: More than one-half of children and adolescents are not learning*. Paris: UNESCO. Retrieved from http://uis.unesco.org/sites/default/files/documents/fs46-more-than-half-children-not-learning-en-2017.pdf

UNICEF (2018). Education Pathway Analysis dashboard. Retrieved from https://data.unicef.org/resources/education-pathway-analysis-dashboard

UNICEF (2019). *A World Ready to Learn: Prioritizing quality early childhood education*. New York: UNICEF. Retrieved from https://data.unicef.org/resources/a-world-ready-to-learn-report

UNICEF (2019). *Survey Findings Report and Snapshot Guidelines*. New York: UNICEF. Retrieved from http://mics.unicef.org/tools#reporting

UNICEF (2019). MICS-EAGLE. New York: UNICEF. Retrieved from https://data.unicef.org/resources/mics-education-analysis-for-global-learning-and-equity

UNICEF (2019). *Do children with disabilities attend school? New findings from Sierra Leone*. New York: UNICEF. Retrieved from https://blogs.unicef.org/evidence-for-action/children-disabilities-attend-school-new-findings-sierra-leone

UNICEF (2019). *ICT skills divide: Are all of today's youth prepared for the digital economy?* New York: UNICEF. Retrieved from https://blogs.unicef.org/evidence-for-action/ict-skills-divide-todays-youth-prepared-digital-economy

UNICEF (2019). *How much do children learn? New evidence from Sierra Leone*. New York: UNICEF. Retrieved from https://blogs.unicef.org/evidence-for-action/much-children-learn-new-evidence-sierra-leone

UNICEF (2019). *Guidelines for Adapting the Foundational Learning Module to Non-Multiple Indicator Cluster Household Surveys*. New York: UNICEF. Retrieved from https://data.unicef.org/resources/guidelines-adapting-foundational-module-non-mics

UNICEF (Forthcoming). *Accuracy Matters: Improving Policy Outcomes by using MICS "School Age" to Calculate Education Indicators*. New York: UNICEF.

UNSTATS (2019). SDG Indicators: Metadata Repository. Retrieved from https://unstats.un.org/sdgs/metadata

UNSTATS (2019). *Indicator 8.7.1 - E-Handbook on SDG Indicators*. UN Statistics Wiki. Retrieved from https://unstats.un.org/wiki/display/SDGeHandbook/Indicator+8.7.1

Washington Group on Disability Statistics (2017). *The Washington Group Short Set on Functioning: Question Specifications*. Washington DC: Washington Group. Retrieved from http://www.washingtongroup-disability.com/wp-content/uploads/2016/12/WG-Document-4-The-Washington-Group-Short-Set-on-Functioning-Question-Specifications.pdf

ANNEX

Examples of SPSS and Stata codes used to prepare education indicators are available in the digital annex. Please visit https://github.com/micseagle/STATA and https://github.com/micseagle/SPSS.

PHOTOGRAPHY CREDITS

Page 12: © UNICEF/UNI232344/Noorani
Page 15: © UNICEF/UNI301193/Stachowske
Page 19: © UNICEF/UN0340017/Frank Dejongh
Page 21: © UNICEF/UN031903/E Ademuyiwa
Page 27: © UNICEF/UN03651/Takyo
Page 37: © UNICEF/UNI156860/Pirozzi
Page 43: © UNICEF/UN038656/Pirozzi
Page 47: © UNICEF/UNI150871/Asselin
Page 48: © UNICEF/UNI284572/Vishwanathan
Page 56-57: © UNICEF/UNI150859/Asselin
Page 59: © UNICEF/UNI180415/Romenzi
Page 61: © UNICEF/UN0251366/Herwig
Page 63: © UNICEF/UN031899/E Ademuyiwa
Page 71: © UNICEF/UN0344474/Prieto

ENDNOTES

1. For more information, please see the 2019 paper *Multiple Indicator Cluster Surveys: Delivering Robust Data on Children and Women across the Globe* at https://onlinelibrary.wiley.com/doi/full/10.1111/sifp.12103.
2. All MICS questionnaires can be found on the MICS website: https://mics.unicef.org/tools.
3. International Telecommunication Union (2014). Manual for Measuring ICT Access and Use. Retrieved from https://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-ITCMEAS-2014-PDF-E.pdf.
4. UNESCO UIS SDG 4 Data Book: Global Education Indicators 2018.
5. https://github.com/micseagle/STATA and https://github.com/micseagle/SPSS.
6. Indicators in grey shading are global indicators, while the remainder are thematic indicators.
7. http://uis.unesco.org/sites/default/files/documents/fs46-more-than-half-children-not-learning-en-2017.pdf.
8. Available SDG4.1.1.i indicators are not globally comparable except for the MICS Foundational Learning module. The projection is based on the countries that signed up for MICS6 and for countries that have not started survey design, assuming that the Foundational Learning and Child Functioning modules are fully taken.
9. https://www.oecd-ilibrary.org/education/isced-2011-operational-manual_9789264228368-en.
10. https://github.com/micseagle/STATA and https://github.com/micseagle/SPSS.
11. Some inconsistencies present in previous versions of MICS should not happen in MICS6 if secondary editing protocols are rigorously followed.
12. The latest set of guidelines available is based on MICS5, but also applies to MICS6. The manual can be downloaded from the MICS website: https://data.unicef.org/wp-content/uploads/2020/02/MICS5_Manual_Processing_the_Data.doc.
13. Each questionnaire has a weight that is identified by the questionnaire code. For example, the questionnaire for children 5–17 is abbreviated fs and the weight in it is fsweight. The other weights are: hhweight (household questionnaire), wmweight (women's questionnaire), mnweight (men's questionnaire) and chweight (children under 5 years old questionnaire).
14. For more information, please see the online FAQ on merging data sets at http://mics.unicef.org/faq.
15. https://github.com/micseagle/STATA and https://github.com/micseagle/SPSS.
16. https://data.unicef.org/resources/mics-education-analysis-for-global-learning-and-equity.
17. For a complete list of module abbreviations, please refer to Figure 3, Chapter 1. CL: Child Labour, EC: Early Childhood Development, ED: Education, FL: Foundational Learning, MA/MMA: Marriage/Union, MT/MMT: Mass Media and ICT, PR: Parental Involvement, UB: Child's Background.
18. SDG4.1.1 calculates the percentage of Grade 2/3 children who successfully completed foundational reading/numeracy tasks.
19. Children enrolled or attending pre-primary education are considered "in school" per the agreement reached in a 2018 meeting of the Technical Cooperation Group on the Indicators for SDG4 – Education 2030.
20. Note that unlike transition rate, the effective transition rate excludes repeaters from the denominator of the formula.
21. The Foundational Learning module is in the process of being updated, so some of the variable names referred to in this text might change.
22. Interestingly, all provinces yield higher odds-ratios of attending school than the omitted province (Vientiane, the capital), which would otherwise be assumed the most privileged. This is due to the fact that most of the reasons children in Vientiane are more likely to be in school (being wealthier and more urban, for example) are already accounted for by the controls in the regression.
23. Age and age squared were omitted from the table. Non-significant values (variables that do not significantly impact attendance) are left blank. Positive values mean that the variable increases attendance while negative values mean they decrease attendance. The base category (the one each variable is compared to) is: male, urban, from the capital province, Lao-Tai and higher-educated mother.
24. The Washington Group was commissioned by the United Nations and mandated to develop comparable measures for disability. The main purpose of the Washington Group is the promotion and coordination of international cooperation in the area of health statistics, focusing on disability measures suitable for censuses and national surveys. The group has developed and endorsed a short set of questions on measuring disability and functioning in the adult population, to be used in censuses and surveys. It has also developed and endorsed an extended set of questions to be used as components of population surveys, as supplements to surveys or as the core of a disability survey that expands on the short set. http://www.washingtongroup-disability.com/wp-content/uploads/2016/12/WG-Document-4-The-Washington-Group-Short-Set-on-Functioning-Question-Specifications.pdf
25. These revisions were motivated by a desire to develop tools that are in line with the WHO International Classification of Functioning, Disability and Health – Children and Youth version (ICF-CY) and the UN Convention on the Rights of Persons with Disabilities. The goal is to assess child functioning in light of barriers and supports to daily living and social participation and to ensure that the entire age spectrum and additional relevant domains are captured.

unicef

for every child

Published by United Nations Children's Fund
Division of Analysis, Planning and Monitoring
3 United Nations Plaza, New York, NY 10017, USA