Achieving Equitable Education

NORRAG SERIES ON INTERNATIONAL EDUCATION AND DEVELOPMENT

Series Editors: Moira Faul and Chanwoong Baek, *Network for International Policies and Cooperation in Education and Training (NORRAG)*

The NORRAG Series on International Education and Development reflects the mission of the global Network for International Policies and Cooperation in Education and Training (NORRAG) to produce, disseminate and broker critical knowledge on topical issues that emerge in education and development. Through its programs, knowledge production and dissemination, NORRAG contributes to enhancing the conditions for participatory, informed and evidence-based policy decisions that improve equity and quality of education. This series aims to act as a knowledge broker at the interface between research, analysis, policy and practice within the comparative, development and international education community.

For a full list of Edward Elgar published titles, including the titles in this series, visit our website at www.e-elgar.com.

Achieving Equitable Education

Missing Education Data and the SDG 4 Data Regime

Edited by

Marcos Delprato

University of Sussex, UK

Daniel D. Shephard

NORRAG, Geneva Graduate Institute and Indiana University Bloomington, USA

NORRAG SERIES ON INTERNATIONAL EDUCATION AND DEVELOPMENT

Edward Elgar PUBLISHING

Cheltenham, UK • Northampton, MA, USA

Cover image: Designed by Anouk Pasquier (source file from starline / Freepik)

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2023949638

This book is available electronically in the Elgaronline
Political Science and Public Policy subject collection
http://dx.doi.org/10.4337/9781035313839

ISBN 978 1 0353 1382 2 (cased)
ISBN 978 1 0353 1383 9 (eBook)

Contents

Contributors

Angela Arnott, born and educated in Zambia, with post-primary schooling in Zimbabwe and tertiary in South Africa, is an education policy analyst with decades of experience working in Africa on building effective education management information systems for policy support and assessments at national, regional, and continental levels. She has worked for SADC, UIS, ADEA, and the Global e-Schools Initiative and is currently supporting the KIX 19 and 21 Hub GPE partner countries with data strategies and research.

Gawaian Bodkin-Andrews is a D'harawal man born in Sydney and is directly connected to (and raised on) Bidigal (Bitter-Water) and Nattaimattagal (Sweet-Water) Country. He is currently the Director of Indigenous Research at Western Sydney University, where his work focuses on Indigenous Studies, Indigenous Research Methodologies, Indigenous Storywork, Indigenous Data Sovereignty, Indigenous education, Indigenous psychology, and racism. At present, his primary research focus centres on leading an Australian Research Council Discovery Indigenous Grant titled *Shielding Our Futures: Storytelling with Ancestral and Living Knowledges.* He is also a member of *Maiam nayri Wingara* the Australia Indigenous Data Sovereignty Collective.

Marcos Delprato is Senior Lecturer at the Centre for International Education, University of Sussex. His expertise lies in the application of quantitative methods in the field of international education and development, focusing on issues related to educational disadvantages in Latin America and sub-Saharan Africa. His empirical research highlights educational inequalities related to early marriage and gender, the impacts of household wealth and school violence on learning, Indigenous discrimination in education systems, private-public school learning gaps, efficiency analysis of secondary education systems, early warning systems to prevent dropouts, and spatial education inequalities.

Karma El Hassan (PhD) is Professor of Educational Psychology Measurement and Evaluation at the American University of Beirut, Lebanon. Up to 1 January 2021 she was Director of the Office of Institutional Research & Assessment (OIRA) at the American University of Beirut. She has extensive experience in teaching undergraduate and graduate courses, training, and research in her area of expertise. In addition, she has conducted and supervised research in test development/adaptation, validation, and use, as well as teacher assessment competencies and published in international and regional journals and books in these areas.

Helen Longlands is Lecturer in Education and International Development at the IOE, University College London's Faculty of Education and Society, and Programme Leader for the MA in Education, Gender and International Development. Her research interests are interdisciplinary and concerned with gender, inequalities and social justice, with a particular focus on masculinities, transnational relationships of power, and the interconnections between the spaces of education, gender, work and family.

Karen Martin is a Noonuccal, Quandamoopah and Bidjara woman. She has taught in all sectors of education where she focuses on early childhood education, Aboriginal education, and Aboriginal research. Her career also includes Ministerial appointments (Australian Government) providing expert advice on Aboriginal affairs, education (systems, professions), and research. Karen is internationally known for her work on decolonising western research paradigms, methodologies, and ethics and applies this expertise as Acting Chair: Longitudinal Study of Indigenous Children (Australian Government Department of Social Services) where her involvement spans almost 20 years. Additionally, this expertise was recognised when she was the only Aboriginal educator invited to the series of Closing the Gap Data Refresh Roundtables (Australian Government, 2018). She is a member of the *Maiam nayri Wingara*, the Australian Indigenous Data Sovereignty Collective where she contributes to activities regarding knowledge translation and education.

Jacob Prehn is a proud Worimi man, Indigenous Fellow-Senior Lecturer, and the Master of Social Work Course Coordinator in the School of Social Sciences at the University of Tasmania. He is an award-winning Early Career Researcher in the fields of sociology and social work and current Co-Investigator on a $1.25million National Health and Medical

Research Council (NMHRC) research grant. He is also a member of *Maiam nayri Wingara*, the Australia Indigenous Data Sovereignty Collective. To date, Jacob's publications cover a range of topics such as Indigenous data sovereignty, Indigenous men and masculinities, Indigenous health and wellbeing, Indigenous education (primary, secondary, and tertiary), and strengths-based approaches.

Martín Scasso has a degree in Sociology from the University of Buenos Aires, specialising in the generation, analysis, and use of statistical information in education for the monitoring, management, and evaluation of policies. He has carried out various studies for governments and international organisations on the educational systems of Latin America and the Caribbean. He is currently a consultant of OREALC/UNESCO Santiago and of CECC/SICA.

Daniel Shephard is the Postdoctoral Fellow in International Sustainable Development at Indiana University Bloomington and the lead of NORRAG's Missing Data Project, Geneva Graduate Institute. His current research focuses on the inclusion of displaced learners in public school systems and makes contributions to comparative and international education, forced migration, analytical sociology, and development studies using mixed methods. His research has involved children, youth, teachers, and educational staff in 30 countries. He holds a PhD and a MPhil in Comparative and International Education from Teachers College, Columbia University and a MSc in Evidence Based Social Intervention from the University of Oxford.

James Shoobridge has almost 30 years' experience in international education development, primarily in monitoring and evaluation, the development of education information systems, education planning, education sector analysis, and other forms of analysis and evaluation. James has worked as both a team leader and consultant on a wide range of projects in the areas of education and public sector reform in over 30 countries throughout Africa, Asia, and Central Asia. In recent years he has been heavily involved in evaluating programmes, developing monitoring systems, sector planning activities, and project design in a number of Asian and African countries. He has also worked extensively in the commercial sector having built and managed an information systems company in Australia which supports both the public and private sectors and is now active in ten countries. James was educated in Australia with a BSc (minor computers), units earned in a PhD in

Science, a graduate diploma in Computing and a Master's in Business Information Technology focusing on monitoring using education information systems.

Elaine Unterhalter is Professor of Education and International Development at University College London, and Co-Director of the Centre for Education and International Development (CEID). She has published widely on gender, education, international and national policy and practice, with a special focus on intersecting inequalities, the capability approach, and approaches to change. Recently published books include *Education and International Development: An Introduction* (co-edited with Tristan McCowan) (2021, Bloomsbury) and *Measuring the Unmeasurable in Education* (2020, Routledge).

Chiara Valenti is a Research Associate at IDMC where she focuses on the socioeconomic impacts and severity assessments of internal displacement. Chiara holds an MSc in Gender, Development and Globalization from the London School of Economics. Her thesis explored the phenomenon of gender-based violence in contexts of displacement and detention, using Libya as a case study. Prior to joining IDMC, she worked with an NGO on the professional inclusion of refugee women in Italy in digital and tech jobs. She has also worked with the Italian Institute for International Political Studies organising events and conducting research on topics of migration and security in the broader Mediterranean region alongside the Italian Ministry of Foreign Affairs.

Alejandro Vera has a degree in Economics from the National University of Córdoba and a Master's in Public Policy from the Torcuato Di Tella University. He is a Programme Specialist of the UNESCO Regional Office of Education for Latin America and the Caribbean (OREALC/UNESCO Santiago). Previously, he was a Social Policy Specialist at UNICEF, National Office for Bolivia, and Assistant Programme Specialist at the UNESCO Institute for Statistics, in the regional team for Latin America and the Caribbean.

Ernesto Yáñez has a degree in Economics from the Bolivian Catholic University, where he currently works as a teacher. He has worked as a consultant in various international organisations on issues related to the economics of education and he is currently a consultant of OREALC/UNESCO Santiago.

Louisa Yasukawa was previously a Researcher with IDMC, where she led research on the socioeconomic impacts of internal displacement, with a particular focus on children, youth and education. She was the lead author of IDMC's report "Informing better access to education for IDPs". Louisa is a qualified lawyer and holds an MSc in Refugee and Forced Migration Studies from the University of Oxford. She is currently an Associate Resettlement Officer with UNHCR in Ecuador.

Rosie Peppin Vaughan is Lecturer in Education and International Development at the IOE, University College London's Faculty of Education and Society. Her research focuses on transnational advocacy on girls' and women's education, and also draws on the capability approach and the concept of human development to explore the measurement of educational equality and social justice.

Abbreviations

AGEE	Accountability for Gender Equality in Education
AU	African Union
CESA	Continental Education Strategy for Africa
COAG	Closing the Gap on Indigenous Disadvantage strategy
CRPD	Convention for the Rights of Persons with Disabilities
DHS	Demographic and Health Survey
EAC	East African Community
ECD	Early Childhood Development
ECE	Early Childhood Education
ECCAS	Economic Community of Central African States
ECOWas	Economic Community of West African States
EFA	Education for All
EGMA	Early Grade Maths Assessment
EGRA	Early Grade Reading Assessment
EMIS	Education Management Information Systems
ERCE	Regional Comparative and Explanatory Study
ESCWA	Economic and Social Commission for Western Asia
GEMR	Global Education Monitoring Report
GER	Gross Enrolment Ratio
GIS	Geographic Information Systems
GPE	Global Partnership for Education
ICCS	International Civic and Citizenship Education Study
ICT	Information and Communication Technology
IDP	Internally Displaced People
IEA	International Association for the Evaluation of Educational Achievement

IGAD	Intergovernmental Authority on Development
IPED	Pan-African Institute for Education for Development
ISCED	International Standard Classification system of Education
MDGs	Millennium Development Goals
MICS	Multi Indicator Cluster Survey
NFE	Non-Formal Education
NSO	National Statistics Office
OCAP	Ownership, Control, Access, and Possession principles
OECD	Organisation for Economic Co-operation and Development
OREALC	Regional Bureau of Education for Latin America and the Caribbean
PASEC	Analysis Programme of the CONFEMEN (La Conférence des Ministres de l'Education des pays ayant le français en partage, the Conference of Education Ministers of Countries Using French in Common) Education Systems
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
REC	Regional Economic Community
SADC	Southern African Development Community
SDGs	Sustainable Development Goals
SEACMEQ	The Southern and Eastern Africa Consortium for Monitoring Educational Quality
TIMSS	Trends in International Mathematics and Science Study
TVET	Technical and Vocational Education and Training
UIS	UNESCO Institute for Statistics
UNDESA	United Nations Department of Economic and Social Affairs
UNDRIP	United Nations Declaration on the Rights of Indigenous Peoples
UNESCO	United Nations Educational Scientific and Cultural Organization

UNHCR	United Nations High Commission for Refugees
UNICEF	United Nations Children's Fund
UNSD	United Nations Statistics Division
USAID	United States Agency for International Development

1. Introduction to missing education data and the SDG 4 data regime

Daniel Shephard and Marcos Delprato

MISSING EDUCATION DATA

The chapters of this book provide a nuanced understanding of education data gaps across regions, themes and levels of education systems. It contributes to an understanding of the relationships and disconnects between national, regional and global data needs. It also highlights the barriers that continue to limit the use of data to inform policy aimed at achieving the shared goal of "inclusive and equitable quality education" for all. This book investigates these issues by focusing on missing education data[1] and its use in relation to SDG 4.

The 17 Sustainable Development Goals (SDGs) were agreed by United Nations member states on 25 September 2015 (UNGA, 2015). The fourth of these (SDG 4) is dedicated to education, specifically to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". The underlying targets are expansive and have helped focus the attention of some international and civil society advocates on achieving a broader vision of education by 2030. We have now passed the midpoint of the agenda, and gaps remain both in terms of achieving the substantive targets of this agenda and of having the data to monitor progress or inform policy changes.

[1] The chapters of this book address three categories of missing education data: (1) data that are missing because of reporting in which data exist but are not reported, including unreported disaggregation or entire unreported indicators; (2) data that are missing because of the absence of data collection; and (3) data that are missing because of the monitoring framework, where key drivers of SDG 4 targets exist in underlying datasets, especially household surveys and learning assessments, but are not part of the SDG 4 monitoring framework.

When looking specifically at data gaps and how they intersect with vulnerable groups and sectors of education, many of those who are being left behind reflect, in part, a historical pattern rooted in previous global agendas and their focus in substance or in data systems on primary, public education for minors. Unfortunately, the groups of learners who may need the most support – and perhaps tend to be most neglected – continue to be the groups that are most beset by missing, inappropriate or unused data. These groups often represent non-dominant populations who continue to be politically marginalized within and across countries.

GOALS AND INDICATORS: A BRIEF HISTORY

The current global agenda and education data regime did not arrive in a vacuum. Rather SDG 4 follows a history of over 30 years of international education goals and indicators, with continuous debates about both positive progress and broken promises to achieve the goals. At Jomtien, Thailand, in March 1990, leaders of 155 low- and middle-income countries alongside heads of multilateral development, civil society and (bilateral) aid agencies released the World Declaration on Education for All (EFA) goals (UNESCO, 1990). The preamble of the document recognized that global partners had failed to fulfil the 40-year-old Universal Declaration of Human Rights Article §26 "everyone has a right to education", since more than 100 million children were out-of-school and 960 million adults were illiterate. To coordinate efforts to address this, the Jomtien conference set out a series of ambitious commitments covering:

- Meeting basic learning needs;
- Universalizing access and promoting equity;
- Focusing on learning;
- Broadening the means and scope of basic education (including early childhood education: ECE), alternative non-formal education provision, literacy programmes for youth and adults and informal education);
- Enhancing the learning environment; and
- Strengthening partnerships.

The keen observer will note that the EFA commitments map onto almost all of the SDG 4 goals – including even its list of groups in need of special attention to address educational disparities. At the end of the 1990s, the degree to which these holistic education goals were considered a success

was measured by quite limited data with declarations such as "developing countries as a whole had achieved net enrolment rates in excess of 80 per cent" (UNESCO, 2000, p. 13), ECE "expanded modestly", non-formal education had seen "gradual growth", and "[w]hile illiteracy remain unacceptably high, a measure of progress [had] been achieved".

At Dakar, Senegal, in 2000, leaders of countries, heads of agencies and civil society organizations met and released the Dakar Framework for Action on "Education for All: Meeting our Collective Commitments" which committed to six EFA goals (UNESCO, 2000). To paraphrase, the goals were:

- Expanding and improving ECE especially for the most vulnerable and disadvantaged children;
- Ensuring that all children have access to and complete free and compulsory primary education of good quality;
- Ensure that learning needs of all young people and adults are met through equitable access to appropriate learning and life skills;
- Improving adult literacy by 50% for, and provide equitable access to, basic and continuing education for all adults;
- Eliminating gender disparities in primary and secondary education; and
- Improving every aspect of the quality of education, and ensuring their excellence so that recognized and measurable learning outcomes are achieved by all.

However, despite these holistic Education for All goals, the Millennium Development Goals (MDGs) agenda, negotiated by nation-states, was adopted later that same year, and focused only on primary school access, completion, and gender parity for children and literacy for youth and adults (UNESCO, 2015). This restriction was especially notable when one looks at the type of data that were highlighted and the types of data that were missing in the final reports. The assessment that significant progress had been made towards this goal since 2000 (UNESCO, 2015) was largely determined by increases in education access in selected subsets of countries with: "nearly 184 million children enrolled in pre-primary" (UNESCO, 2015, p. 45); increased adjusted net enrolment ratios "rising at least 20 percentage points from 1999 to 2012 in 17 countries" (UNESCO, 2015, p. 75); increased "participation in lower and upper secondary education" (UNESCO, 2015, p. 109). Yet at the same time, limited data were available on skills where "the rate of illiteracy is likely to have dropped

slightly from 18% in 2000 to 14% in 2015" (UNESCO, 2015, p. 135); gender parity, with "69% of countries with data having achieved or are likely to achieve gender parity by 2015 (UNESCO, 2015, p. 153); and teachers, since although "pupil/teacher ratios declined in about 83% of the 146 countries with data at the primary education level … in one-third of the 91 countries with data for 2012, less than 75% of primary school teachers were trained according to national standards" (UNESCO, 2015, p. 187). Once again, a broad set of goals were committed to, many of them now echoing in the commitments of SDG 4 that now characterize the global education agenda from 2015 to 2030.

In September 2015, the United Nations General Assembly (UNGA) agreed to the Sustainable Development Goal (SDG) agenda for the years from 2015 to 2030 after years of consultations with nation-states, regional organizations and civil society (UNGA, 2015). The agenda includes SDG 4 on education, with a commitment to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (UNGA, 2015, p. 18). As outlined in more detail in Table 1.1, the SDG agenda includes a more expansive set of targets and indicators than did the MDGs, a set that is in closer alignment with the EFA agenda. Importantly, the agenda also explicitly includes all countries, encompassing the Global North and the Global South.

DATA FOR GLOBAL EDUCATION AGENDAS

The continuous extension of deadlines set in Jomtien, Dakar, and now for the SDGs has led a variety of influential practitioners and thinkers to question whether global education commitments have been effective and sufficiently funded (Klees, 2017; Wulff, 2018), and more fundamentally if they are a form of continued neo-colonialism (Hoppers, 2015; Ntihirageza & Ibrahima, 2021). Others have noted that although progress has been made, the nature of data gaps has meant that those most disadvantaged learners – who have been signalled as a priority since 1990 – are often missing in datasets and, therefore, missing from trends of improvement (or stagnation) (Faul, Montjouridès, & Terway, 2021; IDMC, 2019; Johnstone, Schuelka, & Swadek, 2020; Montjouridès, 2013; Olusanya et al., 2021).

These *missing groups* in the data are precisely the groups that countries need to track and provide social interventions to support. In addition to such missing groups, the disconnect between the holistic commitments and goals of EFA (around ECE, youth and adults, skill development and

learning) were very seldom represented systematically in reports tracking the progress of the EFA goals – especially during the overlapping MDG period (2000–15). Instead of tracking all education levels stated in the EFA agenda (for all age groups and learning outcomes), development reports only tracked primary enrolment and literacy in line with the MDG indicators.

A related challenge has been that broad conceptual goals from SDG 4 have not been translated into the narrower indicators used for reporting (King, 2017). As a result, there have been many systematically missing types of data, and the legacy of the MDG data regime continues to be felt into the SDG period.

There are also important gaps in terms of the use of data and the alignment of data to the needs of policymakers and practitioners who can enact improvements to education systems that could actually improve learning (Buckner, Shephard, & Smiley, 2022; Ntihirageza & Ibrahima, 2021). Such *missing uses* of data would include the use of data to inform subregional education policies and intersections of vulnerability among education stakeholders. These missing uses may also belie a more fundamental mismatch between global education data regimes and local needs (Ntihirageza & Ibrahima, 2021).

An awareness of missing education data is not new. It has featured regularly in both the EFA and later Global Education Monitoring Reports. However, two important points are new regarding data and the SDGs. The first is that we are now living through a period of unprecedented data collection through the use of technology, the trend towards the datafication of governance nationally and globally, and the move (in some contexts) towards open data and open governance (Kitchin, 2014; World Bank, 2021). The second is that the SDGs made an explicit commitment to building the capacity of statistical agencies (SDG 17.19) and the disaggregation of data (SDG 17.18). Unfortunately, national statistical offices saw a further 48 per cent cutback in funding from government and donor sources during the COVID-19 pandemic and continued budgetary shortfalls are expected for years to come (UNDESA, 2022).

We have passed the midpoint of the SDG agenda and the current data regime continues to include systematic gaps that generate missing groups, missing types and missing uses in the data. In 2022, the United Nations warned that "serious data gaps persist in SDG monitoring" (UNDESA, 2022, p. 4). These gaps persist despite over 75 years of commitments to the right to education, 30 years of holistic global education targets, increased data production and capacity, reinvigorated commitments to

achieving quality education for all, and reaching the half-way point of the SDG agenda that explicitly aimed to strengthen national data capacities, to say nothing of calls for "transforming education" (UNESCO, 2022).

INTRODUCTION TO SDG 4 DATA

SDG 4 contains ten targets and 44 indicators. The ten targets along with the number of indicators for each are outlined in Table 1.1.

PRIMARY DATA SOURCES FOR SDG 4

There are four primary data sources for SDG 4 and also for educational planning more broadly:

- Data on institutional structures and policies that are provided by government agencies, primarily the Ministry of Education (e.g., data on whether or not basic education is free and compulsory and how that is defined);
- Data from Education Management Information Systems (EMIS);
- Census-based and sample-based educational assessments including national assessments, regional assessments (e.g., ERCE, PASEC, SEACMEQ), and international large-scale assessments (e.g., PISA, TIMSS); and
- Surveys collected by national statistical agencies alone (e.g., Censuses, Time-Use Surveys, Labor Force Surveys) or with the support of international partners (e.g., Multiple Indicator Cluster Surveys: MICS, or Demographic and Health Surveys: DHS).

The latter three are the primary focus of this book and are also the data sources most used to monitor SDG 4, as they are the data sources that can provide insights for specific categories of learners and stakeholders across the education system(s). To limit repetition across chapters, we provide a brief overview of each of these data sources here.

Systematic collection and analysis of data on education systems exists in all countries. Education management information systems (EMIS) are the primary source of data and are used for day-to-day functions and strategic policymaking by education decision makers at the national and/or state level. Although an ideal EMIS would include the entire ecosystem of data that is relevant to education, in reality traditional EMIS most commonly consist of an annual school census that collects

Table 1.1 Overview of SDG 4 targets and indicators

Target	Description	Number of Indicators
4.1	Ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes	8
4.2	Ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education	5
4.3	Ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university	3
4.4	Substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship	3
4.5	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	6
4.6	Ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy	2
4.7	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	5
4a	Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all	4
4b	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	1
4c	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	7

Note: These are families of indicators that might have subcomponents. For example, 4.1.1 has multiple subcomponents covering different subjects (reading and mathematics), different levels (grades 2 or 3, end of primary, end of lower secondary), as well as disaggregation (e.g., by sex). The number of indicators is based on the UIS Official List

of SDG 4 Indicators published in September 2023. This table excludes targets outside of SDG 4 such as the Education 2030 Framework for Action indicator on government expenditure on education and SDG 1.a on resource mobilization.

data about the public, government-managed, formal school system. More recently, countries with sufficient resources have moved to the use of individual student identifiers and real-time electronic EMIS that update immediately (e.g., when students enrol in or drop out of school, or take an exam). In some countries, data systems can be linked (often via unique identifiers at the individual level) so that EMIS can draw on information from other sectors and government agencies (e.g., health, family services or emergency services). In other countries, EMIS is fragmented and even different directorates of education are not interoperable. Thus, EMIS span from real-time, individual, interconnected data systems to annual, aggregate, disconnected data systems. While some EMIS integrate granular data on both "outputs" (e.g., schools, enrolment) and "inputs" (e.g., budgets, purchases, teacher salaries, etc.), others do not integrate these even though "input"-related data might provide useful information on key drivers of educational outcomes, such as teacher profiles, supports, benefits and turnover.

Assessment data might be census-based (e.g., all learners at a certain grade level) or sample-based (e.g., with a nationally or subnationally representative sample of learners). Some assessments are aligned with the curriculum (e.g., most national assessments and – to a lesser degree – some cross-national assessments such as TIMSS) while others are based on broader standards of proficiency in a certain domain (e.g., PISA). Although challenges remain with aligning assessments to a common proficiency framework, there have been important advancements in this area. In 2019, the Global Proficiency Framework for Reading and Mathematics was released that outlines expected skills for reading and mathematics at the relevant grade levels for reporting on SDG 4.1.1 (UNESCO, 2019). This has been accompanied by a methodology for mapping existing assessments to the global proficiency framework. However, there is much less clarity on how to align measurements related to SDG 4.7 and more holistic educational outcomes including education for sustainable development and global citizenship, although there are some proposals for using existing data to do so (Sandoval-Hernández, Isac, & Miranda, 2019). Assessment data are collected with varying regularity, ranging from annually to every few years. Some assessments include information on important drivers of assessment outcomes (e.g., students' and families' socio-economic status) while others do not.

Assessments range in cost and complexity, but usually require a substantial investment.

Surveys can provide the most detailed information on certain aspects of education, especially education activities and skill levels that are outside of the formal, government-managed education system. Surveys can also enable the linking of important drivers of educational quality and equity – such as socio-economic status and social norms. However, surveys are infrequent, costly, and may not have sufficient sample sizes or sampling frames to capture important vulnerable groups. Examples of surveys that inform education data range from rare full censuses (which are used in education statistics for determining the reference category for enrolment and out-of-school rates) to regular, sample-based population surveys (e.g., Labour Force Surveys that might include information on technical and vocational education participation). In many low- and low-middle-income countries, irregular surveys supported by international donors are relied on for a number of education statistics (e.g., DHS and MICS).

The global education data regime is not simply a technocratic machine, it is both shaped by politics and interests, and in turn exerts influence on global and national education systems (Fontdevila & Grek, 2020; Grek, 2022; Hoppers, 2015; King, 2017; Olusanya et al., 2021; Unterhalter, 2019). Data are both a manifestation of our life worlds and an instrument that shapes them, and in so doing can perpetuate and deepen inequalities (Hoppers, 2015; Mignolo, 2011; Ntihirageza & Ibrahima, 2021). Therefore, throughout this book the authors have addressed not just the technical, but the social and political dimensions that influence (and are influenced by) data production and use.

OVERVIEW OF BOOK CHAPTERS

The book is organized into eight substantive chapter contributions followed by a concluding chapter. In Chapter 2, Marcos Delprato provides a conceptual framework for considering missing education data along with an argument for the more systematic use of data in a way that aligns with the education literature on the factors that matter the most for access to and achievement in education.

Chapters 3 through 6 then provide four regional perspectives on the gaps in terms of missing groups, types and uses of education data. In Chapter 3, Alejandro Vera, Ernesto Yáñez Aguilar and Martín Guillermo Scasso discuss data gaps in – and implications for – education in Latin

America and the Caribbean, with a focus on learners on the move and those with disabilities. They contrast the drop in global data reporting with the increased data production in the region and make a compelling argument for the need to better support countries to link and harmonize their data systems internally and regionally. In Chapter 4, James Shoobridge argues that the education – and education data – challenges facing Asia are reflective of global difficulties, including challenges of data disaggregation, learning outcomes and education outside of the formal public system. He focuses our attention on the data gaps that continue to leave behind learners in conflict regions and those attending non-government schools. In Chapter 5, Karma El Hassan focuses on the data gaps across the Arab States, highlighting the extreme diversity of countries in the region and the challenges of cross-country learning and alignment of education priorities and data systems. She calls for a renewed focus on education data that can support displaced persons, low-income children, and persons with disabilities, pointing to continued gaps in terms of data on educational affordability, outcomes "of relevance for social and civic life" and early childhood education. In Chapter 6, Angela Arnott argues that there is a need to improve the alignment of global targets and data demands with national and regional priorities and resources in Africa. She calls for a renewed focus on addressing priority education data gaps regarding early childhood education, youth and adult learners, learning achievement and skills.

Each of the following three chapters covers specific thematic topics that are particularly relevant for exploring missing education data and the connections between data and achieving quality education for all. The thematic areas covered are internal displacement, indigeneity and gender. In Chapter 7, Chiara Valentini and Luisa Yasukawa provide a compelling argument for the urgency of understanding and addressing the vulnerabilities faced by internally displaced learners as an integral part of achieving SDG 4. The authors point to emerging promising practices that can help address the gaps in current education data systems in order to help policymakers better meet the educational needs of these learners. In Chapter 8, Jacob Prehn, Karen Martin and Gawaian Bodkin-Andrews provide a strong argument that in order to address systemic gaps in education data and provision, Indigenous people's data sovereignty and data governance must become the norm. Their chapter reminds us of the complexities and historical injustices embedded in the global education regime's unitary conceptualization of "nation-states". The chapter demonstrates that the United Nations Declaration on the Rights of Indigenous Peoples (UNGA,

2007) provides stronger support for Indigenous educational rights and data sovereignty than does the SDG 4, which limits its characterization of Indigenous peoples to a vulnerable group whose data need to be disaggregated. In Chapter 9, Helen Longlands, Rosie Peppin Vaughan and Elaine Unterhalter provide a critical analysis of the intersection of education, data and gender equity, and the work being done to move from simplistic notions of gender parity to a more holistic, capability-based conception of gender equity in education as well as the data that would help reflect and inform that.

The book concludes with reflections on common themes across the chapters and, most importantly, some of the shared implications across the contributions for the ways in which changes to the global education data regime can better contribute to reaching the goal of quality, equitable, education for all.

REFERENCES

Buckner, E., Shephard, D.D., & Smiley, A. (2022). Beyond numbers: The use and usefulness of data for education in emergencies. *Journal on Education in Emergencies*, *8*(1). doi:https://doi.org/10.33682/tgfd-m9eg

Faul, M.V., Montjourides, P., & Terway, A. (2021). Education and the new inequality divides. *Global Challenges*, 9.

Fontdevila, C., & Grek, S. (2020). The construction of SDG4: Reconciling democratic imperatives and technocratic expertise in the making of global education data? In *World yearbook of education 2021* (pp. 43–58). New York: Routledge.

Grek, S. (2022). The education Sustainable Development Goal and the generative power of failing metrics. *Policy and Society*, *41*(4), 445–57. doi:10.1093/polsoc/puac020

Hoppers, C.A.O. (2015). Think piece: Cognitive justice and integration without duress. The future of development education – perspectives from the South. *International Journal of Development Education and Global Learning*, *7*(2), 89–106.

IDMC. (2019). *Twice invisible: Accounting for internally displaced children*. Geneva: IDMC.

Johnstone, C.J., Schuelka, M.J., & Swadek, G. (2020). Quality education for all? The promises and limitations of the SDG framework for inclusive education and students with disabilities. In A. Wulff (Ed.), *Grading Goal Four* (pp. 96–115). Leiden, the Netherlands: Brill and Sense.

King, K. (2017). Lost in translation? The challenge of translating the global education goal and targets into global indicators. *Compare: A Journal of Comparative and International Education*, *47*(6), 801–17. doi:10.1080/03057925.2017.1339263

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. London: Sage.
Klees, S.J. (2017). Will we achieve Education for All and the education Sustainable Development Goal? *Comparative Education Review*, *61*(2), 425–40. doi:10.1086/691193
Mignolo, W.D. (2011). Geopolitics of sensing and knowing: On (de)coloniality, border thinking and epistimic disobedience. *Postcolonial Studies*, *14*(3), 273–83.
Montjouridès, P. (2013). Education data in conflict-affected countries: The fifth failure? *PROSPECTS, 43*(1), 85–105.
Ntihirageza, J., & Ibrahima, A. (2021). Education and epistemicide in Africa: Towards an ubuntu-based comprehensive model of education. In R. Tshifhumulo & T.J. Makhanikhe (Eds.), *Handbook of research on protecting and managing global Indigenous knowledge systems* (pp. 275–94). Hershey, PA: IGI Global.
Olusanya, B.O., Hadders-Algra, M., Breinbauer, C., Williams, A.N., Newton, C.R.J., & Davis, A.C. (2021). The conundrum of a global tool for early childhood development to monitor SDG indicator 4.2.1. *The Lancet Global Health*, *9*(5), e586–e587. doi:10.1016/s2214-109x(21)00030-9
Sandoval-Hernández, A., Isac, M.M., & Miranda, D. (2019). *Measurement strategy for SDG global indicator 4.7.1 and thematic indicators 4.7.4 and 4.7.5 using international large-scale assessments in education*. Paris: UNESCO.
UNDESA. (2022). *The Sustainable Development Goals report*. New York: United Nations Department of Economic and Social Affairs (UNDESA).
UNESCO. (1990). *The World Declaration on Education for All*. Paris: UNESCO.
UNESCO. (2000). *The Dakar Framework for Action – Education for All: Meeting Our Collective Commitments*. Paris: UNESCO.
UNESCO. (2015). *2015 Global Monitoring Report – Education for All 2000–2015: Achievements and challenges*. Paris: UNESCO.
UNESCO. (2019). *Global Proficiency Framework for Reading and Mathematics*. Paris: UNESCO.
UNESCO. (2022). Driving a global movement to transform education: Key moments of 2022. https://www.unesco.org/en/articles/driving-global-movement-transform-education-key-moments-2022 accessed 6 September 2023.
UNGA. (2007). United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP). United Nations.
UNGA. (2015). Transforming our world: The 2030 Agenda for Sustainable Development. In *Resolutions of the General Assembly* (Vol. A/RES/70/1). New York: United Nations.
Unterhalter, E. (2019). The many meanings of quality education: Politics of targets and indicators in SDG4. *Global Policy*, *10*(S1), 39–51. doi:https://doi.org/10.1111/1758-5899.12591
World Bank. (2021). *Data for better lives*. Washington, DC: World Bank.
Wulff, A. (2018). SDG 4: The quest for public funding for education and SDG 4. In B. Adams, R. Bissio, C.Y. Ling, K. Donald, J. Martens, S. Prato, & S.

Vermuyten (Eds.), *Spotlight on Sustainable Development 2018: Exploring new policy pathways* (pp. 116–18). Bonn: Druckerei Brandt GmbH.

2. A conceptual framework to assess missing data for SDG 4

Marcos Delprato

INTRODUCTION

Less than a decade remains before the SDG agenda's deadline in 2030. Yet achieving SGD 4 seems quite uncertain. Indeed, latest global estimates for access are quite alarming with around 258 million children, adolescents and youth (which is 17 per cent of the global total) not in school (UNESCO, 2020) and, if current trends persist, only six out of ten young people will complete secondary education by the SDG deadline (UIS, 2019). There is also a need to boost quality education, as between 55 and 60 per cent of students at primary and lower secondary school levels are not reaching minimum proficiency levels in reading and maths (UIS, 2019). According to the World Bank (2019), by the 2030 deadline, about 43 per cent of late-primary children in low- and middle-income countries will still not have reached minimum proficiency in reading by 2030. The potential for monitoring of within-country educational inequalities to ensure that no learners are left behind is also hampered, since around 40 per cent of countries lack disaggregated data on key education indicators from household surveys (UNESCO, 2020).

According to UNESCO (2021): "data gaps on key SDG4 indicators remain a major bottleneck to tracking and monitoring progress" (p. 30). This chapter's analysis departs from recent reports in two ways. First, it employs a novel framework to analyse missing data along different dimensions that have the potential to more precisely inform educational planning and policy to support learners' attainment and learning outcomes, especially those who are being left behind. Second, compared to UIS (2019, 2020), the analysis goes beyond patterns of single missing data points from the 2010–20 period, showing how the lack of informa-

tion can also depend on a country's level of income and whether it has been affected by conflict.

Thus, the framework I offer incorporates relevant information provided from household and learning surveys, and identifies the concrete and significant intersections between the drivers of attainment and learning inequalities beyond SDG 4 indicators. Specifically, the framework proposed is a more efficient mechanism for monitoring SDG 4 targets because it uses broad drivers and their intersections, where these intersections are constructed by concentric layers of disadvantages that shape the results shown in the education indicator.

MEASURING LACK OF PROGRESS TOWARDS SDG 4

The reported lack of progress in SDG 4 indicators for access, attainment, learning and equity calls for more refined, accurate and useful data on education, as well as an assessment of which obstacles are impeding significant progress on SDG 4. A key obstacle is missing data, namely, which information could be possibly included within SDG 4 monitoring platforms in order to better identify and effectively target sub-groups, especially those groups who experience compounded layers of disadvantages and are left furthest behind. Broadly, one could think of three categories of "missingness" in relationship to SDG 4 indicators:

- *Missing due to reporting* (MR). Even though the data exist, they are not reported in the UIS data[1] (perhaps due to quality and comparability issues) and could represent either a missing disaggregation or a missing whole indicator.
- *Missing due to absence* (MA). In this case, absence is due to the lack of data collection.
- *Missing due to the framework* (MF). Albeit important for achieving SDG 4, key drivers of SDG 4 targets which are available from household surveys and learning assessments remain absent from the UIS database framework.

Because the basis of comparison is the UIS database, which is mandated to compile SDG 4 indicators, our focus in this chapter is on the MR and MF categories of missingness.

---

[1] See http://data.uis.unesco.org

Table 2.1 Monitoring SDG 4: missingness against dimensions

		Dimensions	
		1D	2D+
Missingness	Missing reporting (MR)	MR-data1D	MR-data2D+
	Missing framework (MF)	MF-data1D	MF-data2D+

In addition, there is the dimensional aspect of missing data linked to how education indicators can be disaggregated by looking at the intersection of multiple indicators. There could be lack of data for one dimension or one sub-group as a breakdown for the indicator, or two (and more) dimensions' breakdowns obtained by combining dimensions and layers of inequality for the indicators. Taking into account such dimensionality is necessary to enable data to inform education planning for those who are being left behind.

The approach proposed for uncovering missingness in existing data is an intersection of either one or more dimensions of an indicator and whether the indicator or its dimensions are missing due to reporting (MR) or the framework (MF). Table 2.1 illustrates this approach for applying the framework resulting from the consideration of the *categories* of missing data against the *number of dimensions* (one or two) for which it is missing in the UIS database. First, the data that are missing due to reporting (MR) could be missing for one-dimension (1D) from the UIS database (MR-data1D), because it is currently used for another indicator, while MR-data2D denotes a two-dimensional (or more) (2D+) combination using dimensions available in the UIS data. Second, there is the missing data due to the framework (MF), where leading drivers behind educational inequality are not used or included in the UIS data, but they are available in household and learning surveys (so, they are "new" dimensions considered). Here, missing MF-data1D represents a single disaggregation dimension (1D) that could be derived from the underlying data source, whereas MF-data2D denotes a two-dimensional (or more) (2D+) combination of these new single dimensions with themselves.

With this foundation in place, we can now consider how to integrate relevant data from beyond UIS data to give a more detailed and intersectional view of the drivers of educational attainment and learning at the levels of the learner, their household, and school and community.

FURTHER CONCEPTUAL FRAMEWORK FOR LEARNING AND ATTAINMENT

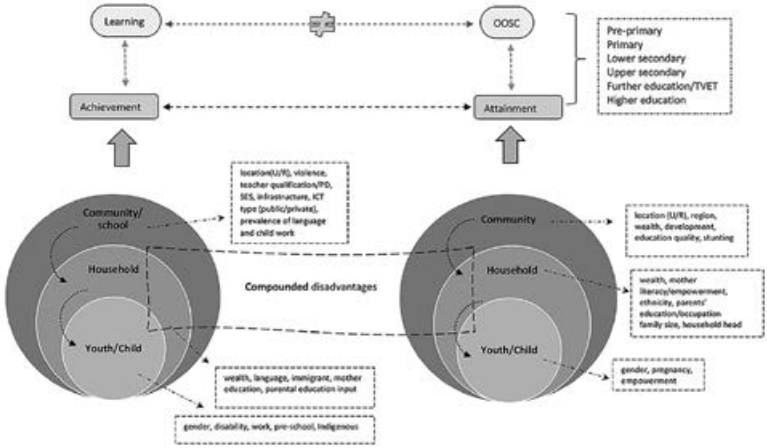
In this section I introduce a more granular assessment of education inequality behind SDG 4 targets, focusing on learning and attainment indicators. This heuristic framework is a representation of layered sources of inequalities, starting at the child/youth level and then moving-up to their households, schools and communities. The framework is more comprehensive than UIS data as it entails:

> (i) a life-course approach to assess student achievement and their drivers, being more relevant to resource-constrained education systems with larger populations of marginalised students (Willms, 2018), also placing an emphasis on a country education sub-systems (e.g., school types); (ii) economic constraints at the household level, by composition and size (demographic factors) and parental education (as mechanism to proxy intergenerational inequalities); (iii) cross-over with SDG5 (gender equality) through measures of women empowerment at the household or community levels (social norms); (iv) an incorporation of the effects of all drivers at various levels simultaneously as compounding factors. (e.g., Huisman & Smits, 2015)

The suggested framework (Figure 2.1) is pragmatic and can be used by those working with detailed education datasets (e.g., national monitoring organisations, academics and decision makers related to SGD 4). In particular, the framework, which relies on current available information from learning and household surveys, is a more efficient mechanism for monitoring SDG 4 targets because it uses broad drivers and their intersections, where these intersections are constructed by concentric layers of disadvantages that shape the results shown in the education indicator.

Learning achievement indicators (on the left of Figure 2.1) are incomplete if they do not take into account the foundations of education systems that help students succeed in their school life; that is, inclusive environments, learning time and quality instruction, family/community support and material resources (Willms, 2018).[2] I operationalise these using nested categories of school and community, household and individual child/youth. On the outermost layer, I use contextual information such as school location and type (urban/rural; public/private), school disadvantage (measured by average family socio-economic status (SES), the

[2] Willms's (2018) framework for learning indicators is focused on primary and secondary levels; this chapter goes beyond that.



Notes:
(1) Learning: based on the educational prosperity framework behind PISA-D (OECD, 2018; Willms, 2018).
(2) Attainment drivers (e.g., completion, out-of-school rates) follows empirical research (e.g., Huisman & Smits, 2015). Acronyms definitions: OOSC, out-of-school children); attainment, completion of different education levels; location, U = urban, R = rural; PD, teachers' professional development; SES, socio-economic status; prevalence language, proportion not speaking the language of the test; parental education input (help of parents with schoolwork of their children; community development, index capturing housing quality, etc.; education quality, community literacy rate; empowerment, women empowerment measured by early marriage, empowerment on mobility, economic decisions, etc.; household head, present in the household.

Figure 2.1 Monitoring – heuristic framework

prevalence of students working outside school, etc.), and schooling processes (indicated by the proportion of qualified teachers and a school's physical and ICT infrastructure). At the family level, the framework uses standard learning drivers such as language spoken at home, household wealth, parental education and parental involvement in their children's education. At the level of the individual student, the standard drivers include gender, disability, Indigeneity, etc. Additionally, multifaceted inequalities also come from interacting community or school factors that overlap with student and family factors; for instance, poor rural Indigenous girls, poor rural students attending public schools, etc. The new dimensions – and their combinations – which I propose for monitoring progress on SDG 4, follow this logic of measuring educational and

non-educational drivers of education outcomes as well as the ways in which these overlap and intersect.

For attainment and access indicators (shown on the right of Figure 2.1), the framework divides the context into individual, household and community as concentric levels.[3] Some factors are applicable to indicators happening later in the life course, or are gendered (e.g., lack of empowerment through early marriage or pregnancy) and therefore linked to SDG 5. Other factors include decisions on educational participation made at the household level by parents, which, in turn, are influenced by poverty, family composition and parental literacy. Moreover, analyses of community location overlapped with subnational regions can serve as a proxy for school supply, while stunting rates in a community provide information on barriers for educational investment (returns to schooling) from parents. In short, by using information available for existing datasets in novel ways, the framework offers a better profiling of which intersectional groups are the most disadvantaged within education systems, and so it lowers the chances that specific disadvantaged and hard-to-reach groups are being excluded from the monitoring framework of SDG 4.

Approach for Assessing Missing Data

As an exercise to identify missing data of different kinds, one could follow the following three steps:

1. Check data sources for each SDG 4 indicator from the official monitoring platform UIS data and, if the indicators employ either household or learning surveys, check for missing dimensions (go to step 2).
2. If there are missing combinations across the existing monitoring dimensions (missing because of reporting: **MR**), include them. They consist of combinations: dimension 1 (*MR-data1D*), with dimension 2 plus dimension 3. This is denoted as: *MR-data2D+* (or missing because of reporting due to missing combinations).
3. Check existing dimensions used in the UIS data for the given indicator against the heuristic monitoring framework proposed (Figure 2.1).

[3] Community is defined by the sample unit available in household surveys (e.g., DHS, MICS) which can be used as an indirect indicator for contextual poverty and development, and also the quality of education where children and youth live.

If these dimensions are missing from the organising principles of the UIS databases, this is called missing because of framework: **MF**.

- If there are dimensions missing from the UIS data, include them (as dimension 1): *MF-data1D*.
- Overlap different new dimensions, generating two- and three-dimension overlaps. This is denoted as: *MF-data2D+* (i.e., *MF-data1D-i* by *MF-data1D-ii* = *MF-data2D+-i*).

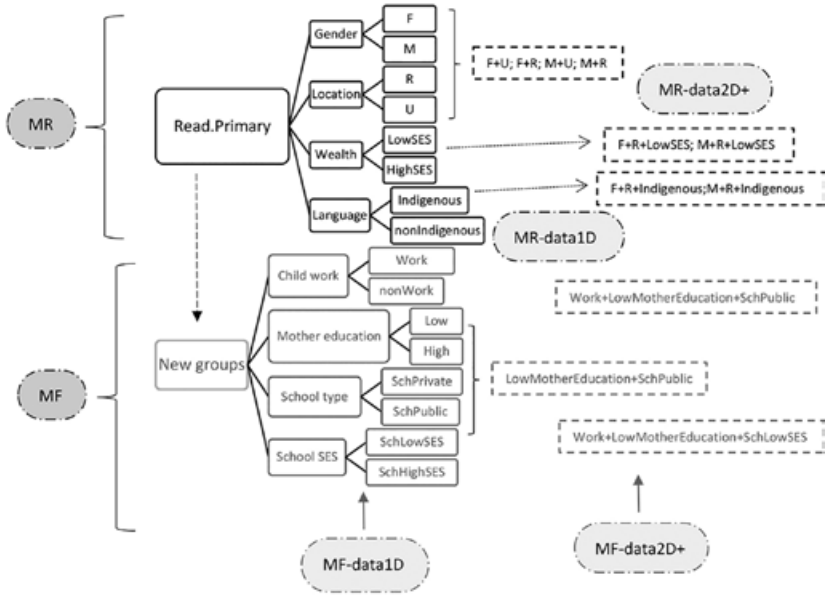
Below, I show a worked example for missing data for SDG indicator 4.1.1 outcome "reading at the end of primary" (Read.Primary). Figure 2.2 shows the existing monitoring categories in the UIS data of disaggregation for the Read.Primary indicator, which are: *gender* (male: M, female: F), *location* (urban: U, rural: R), *family wealth* (low: LowSES, high: HighSES), and whether the child speaks the *language of the test* (L1Yes, L1No), which is a proxy indicator of Indigeneity.[4]

First, the UIS data do not provide analysis of interactions across these four dimensions (missing because of reporting: MR). Thus, *MR-data2D+* can be generated by adding new interactions across them. For instance, interacting gender and location: F+U;F+R;M+U;M+R (2 dimensions) and, further, interacting these two features with either wealth or language will generate a new three-dimension overlap across categories: F+R+LowSES;M+R+LowSES and F+R+L1Yes;M+R+L1No.[5]

Second, some missing groups (MF) can be identified because the framework I describe (*MF-data1D*) measures intersections between poverty, educational transfer across generations, and differential on schooling process and segregation, including: child work (Work, nonWork), mother education (LowMotherEducation, HighMotherEducation), school type (SchPrivate, SchPublic) and school average family SES (SchLowSES, SchHighSES). When these newly

[4] Throughout the chapter, I denote monitoring category overlaps by three elements: using the prefix of each category, the plus sign to indicate the combination of two (or more) categories and, finally, a box (in the figures only) to encircle all combinations possible stemming from the initial categories. I recognise the numerous challenges of the commonly used language proxy for Indigenous peoples, the challenges associated with such proxies and the governances of Indigenous data are explored in more depth in Chapter 8 of this volume.

[5] Also, four dimensions can be obtained by adding family wealth (e.g., poor-rural-Indigenous-girls). Note, though, that in some combinations and surveys, it might not be possible to have a large enough sample ($N \geq 30$) to allow for a reliable estimate of intersected dimensions.



Notes:
(1) Author's constructed example showing existing monitoring categories (in black) and proposed new available missing groups (in grey).
(2) The figure also shows the four types of missingness due to reporting and due to framework.

Figure 2.2 Assessing missing data example: Read.Primary (SDG indicator 4.1.1)

identified groups are interacted (i.e., *MF-data2D+*), further marginalised groups are identified. For instance, two-dimension intersections, such as LowMotherEducation+SchPublic or three-dimension overlaps, for example, Work+LowMotherEducation+SchPublic, are effective sub-categories for monitoring drivers of progress towards SDG 4. In summary, the example above shows that by checking which drivers behind the outcome Read.Primary are missing, one could assess additional crucial layers of disadvantage which are omitted in the current monitoring frameworks of SDG 4.

MISSING DATA ANALYSIS

In this section, I present some patterns of missing data from household and learning surveys that are used in SDG 4 indicators available in the UIS data but where their dimensions are not included by UIS.[6] I show some leading results on the conceptual dimensions excluded across education indicators by targets and types.

New SDG 4 Data Combinations

Table 2.1 shows that, for the 14 SDG indicators (derived from household and learning surveys) of the UIS data platform there are currently 267 breakdowns (1D) or new dimensions for all educational outcomes by which indicators could be disaggregated (column 3). Additionally, it identifies 134 2D breakdowns and 80 breakdowns for three-group overlaps (3D) (columns 4 and 5). There are around 45 per cent (N = 390) more missing dimension combinations in comparison to existing dimensions/drivers, out of which 75 per cent are two-dimension combinations and 25 per cent are three-dimension combinations. Notably, compounded disadvantages are only provided for five (2D) or four (3D) indicators out of the 14 SDG indicators. Nevertheless, there is scope to identify further indicators of inequalities by generating combinations using some existing 1D groups (i.e., the first layer used to identify the sub-groups and the drivers of their education inequalities).

For instance, for learning outcomes (measured by target 4.1.1), the UIS data neither uses 2D nor 3D layers of learning-driven disadvantages,

[6] Relying on these types of surveys might restrict number of countries, so missing rates may be seen as lower bound estimates.

Table 2.2 Missing SDG 4 data disaggregation – missing combinations

SDG 4 target	Non-missing combinations				Missing combinations			Rate
	All	1D	2D	3D	All	2D	3D	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
4.1.1	66	66			312	216	96	473%
4.1.2	102	30	42	30	30	30		29%
4.1.3	6	6						
4.1.4	102	30	42	30	30	30		29%
4.1.5	6	6						
4.2.1	3	3						
4.2.2	44	10	24	10				
4.2.3	7	7						
4.3.2	34	10	14	10	10	10		29%
4.4.3	27	27						
4.5.2	21	21						
4.6.2	27	15	12					
4.a.1	22	22						
4.a.2	14	14			8	8		57%
Total	481	267	134	80	390	294	96	81%

Notes:
(1) Missing combinations based on existing UIS data (http://data.uis.unesco.org).
(2) 1D: one dimension; 2D: two dimensions overlap; 3D: three dimensions overlap.
(4) Rate = [missing combinations (all) / non-missing combinations (all)] / (number of outcomes).

although it is still possible to generate numerous overlaps (= 312) from 2D (= 216) and 3D (= 96) (columns 6 to 8). For completion (indicator 4.1.2) and out-of-school (indicator 4.1.4) rates, even though there are already monitoring sub-dimensions, it is possible to generate around 30 additional identifiable sub-groups (which represent 7–10 per cent of the existing dimensions for these two indicators). Additionally, for indicator 4.a.2 (bullying rates at primary and lower secondary levels), dimensions' overlaps are overlooked, but they can be generated using common drivers of school violence (i.e., household wealth and gender).

Figure 2.3 provides further disaggregation on how missing combinations fluctuate across indicators and also which are the main features for these missing dimension combinations. The exclusion of combinations

affects first (and more strongly) learning indicators and, second, completion/access indicators. In particular, the plot on the left of Figure 2.3 shows that 52 combinations are not included in the UIS data platform for maths and reading achievement, from grades 2–3 up to lower secondary. Completion rates from primary to upper secondary have around ten missing combinations each, the same as for out-of-school rates, while bullying indicators have four missing combinations. The exclusion of these sub-domains clearly impacts on the degree of efficiency when monitoring SDG indicators 4.1.1, 4.1.2 and 4.1.4. Certainly, the omitted type of combinations are quite powerful drivers of education inequality: household wealth, gender and language spoken at home (see right panel of Figure 2.3).

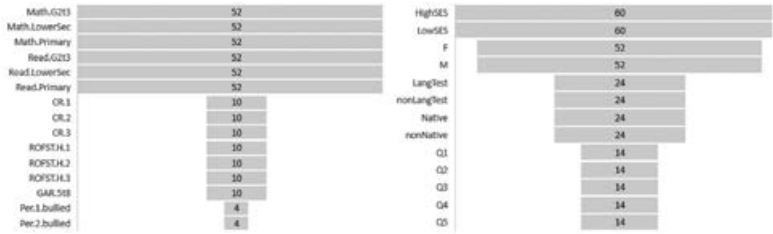


Figure 2.3 *Missing SDG 4 data disaggregation dataset – missing combinations*

Newly Identifiable Groups

Populating SGD 4's monitoring architecture by identifying new, intersectional groups whose education progress (or lack of it) could be monitored would translate into a more nuanced identification of the most marginalised and vulnerable groups and, equally, it would highlight systemic and wider contextual marginalisation. The inclusion of these groups is inexpensive as their data are already available in (or easy to derive from) existing datasets.

Moreover, although this is an ad hoc conceptual exercise, I put forward ideas for some plausible newly identifiable groups (=104) following the heuristic framework (i.e., Figure 2.1). A snapshot of the proposal is described in Table 2.2 (columns 2 and 6 are key: newly identified groups and newly identified group details). Most intersectional drivers identified are applicable to learning and access/completion outcomes. These out-

comes have between 14 and 18 new dimensions (Table 2.2, column 2). I discuss the major findings below.

First, for learning achievement outcomes, across the 14 novel indicator breakdowns,[7] I recommend examining dimensions moving outwards along the concentric circles in Figure 2.1, from initial markers of disadvantage linked to the child/student situation (e.g., child work, disability, preschool attendance) to the family (e.g., mother education, parental education input), and then to schooling conditions (ICT and infrastructure, teacher qualifications) and contextual factors (poverty, school rate of child work, school SES) (column 6). The combination of new intersectional group characteristics with learning outcomes yields 136 novel outcomes to monitor.[8] Importantly, new sub-groups identified through combining intersectional drivers of learning allow education stakeholders and decision makers (in a targeted and scaffolded way) to focus on supporting their learning and targeting barriers that hold back learning for deprived students. Second, for completion (15 new intersectional groups; indicator 4.1.2) and out-of-school outcomes (18 new groups; indicator 4.1.2), the community and region become key units of monitoring to identify how changing factors at these levels can address inequality in education. The choice of new, intersectional groups here is aimed at identifying mechanisms which drive educational outcomes in groups marginalised due to demographic and family composition factors including parental occupation and education, as well as deeper factors such as women's empowerment and prevailing social norms in a community (column 6).

All in all, the newly identified groups for these indicators result in 67 (completion rates) and 107 (access/out of school) new indicators using existing data. Chiefly, here I include 2D-3D interactions across these new groups, linked to an incremental and cumulative geographical lens for informing the improved targeting of groups of learners being left behind. For example, for the case of primary completion rate (CR.1), this results from disaggregating the average country value of the indicator by location (urban/rural), community poverty (low or high) and characteristics

[7] Note that here I don't include 2D and 3D overlaps; this is a methodological choice but, certainly, several overlaps could be incorporated (e.g., for 3D overlaps: child work + low mother education + schoolSES low; disability + low parental input + public school).

[8] The number of new outcomes can be reduced (halved) by using ratios for each new group proposed.

Table 2.3 Missing SDG 4 data disaggregation – new groups

SDG 4 target	Newly identified groups	2D	3D	New outcomes	Newly identified groups details – 1D
(1)	(2)	(3)	(4)	(5)	(6)
4.1.1	14	0	0	136	child work, disability, mother education, parental education input, preschool, school child work, school ICT, school infrastructure school language test, school SES, schoolteacher professional development, schoolteacher qualification, school type, school violence
4.1.2	15	4	1	67	family head, family large, father occupation, girl empowerment early marriage, girl empowerment pregnancy, location mother education, mother empowerment, early marriage, mother empowerment expenditure, mother empowerment mobility mother empowerment reproduction, mother literacy, mother work, region, religion-ethnicity
4.1.3	3	2	1	42	location, region, wealth
4.1.4	18	7	2	120	family head, family large, father occupation, gender, girl empowerment early marriage, girl empowerment pregnancy, location mother age at birth, mother education, mother empowerment early marriage, mother empowerment expenditure mother empowerment mobility, mother empowerment reproduction, mother literacy, mother work, region, religion-ethnicity, wealth
4.1.5	8	3	2	112	family head, family large, father occupation, location, mother education, mother work, region, wealth
4.2.1	7	4	1	62	family head, location, mother empowerment all, mother empowerment early marriage, mother literacy, region, wealth
4.2.2	2	1	1	13	mother age at birth, mother education
4.2.3	3	3	1	25	location, mother education, wealth
4.3.2	4	2	1	13	family large, location, mother empowerment early marriage, region
4.4.3	3	3	2	414	location, region, wealth

SDG 4 target	Newly identified groups	2D	3D	New outcomes	Newly identified groups details – 1D
4.5.2	8	2	1	82	location, school child work, school infrastructure, school SES, schoolteacher prof dev, schoolteacher qualification, school type, wealth
4.6.2	7	4	2	112	
4.a.1	5	2	1	62	school language test, school location, school location size, school SES, school type
4.a.2	7	4	1	594	language test, neighbourhood violence, nuclear family, repeat, school language test, school type, wealth
Total	104	41	17	1854	

Notes:
(1) 2D groups: community development, community education quality, community stunting, community wealth, gender, language test, location, mother empowerment, early marriage, mother literacy, school location, school socio-economic status (SES), wealth.
(2) 3D groups: gender, region, school SES.
(3) Although 2D, 3D groups combinations are included, the focus is on the second column (new groups).

of regions of the country (e.g., if it is rural or urban).[9] Likewise, I propose 3D overlaps for out-of-school rates (ROFST.H), and having region as one domain alongside community location and its development index.

Monitoring Indicators

This section contains an overview on missing patterns for educational outcomes employed to monitor SDG 4 at the country level, assessing whether a given country value indicator is present using the UIS data (14 indicators and 58 outcomes) for the 2010–20 period. I further check missing patterns by wave (wave 1: 2010–15, wave 2: 2016–20) and by country income groups, regions and by the presence of conflict (Table 2.3).[10]

[9] The new indicators for this 3D overlaps are: CR.1.URB.CQ12.R, CR.1.URB.CQ345.R, CR.1.RUR.CQ12.R, CR.1.RUR.CQ345.R.

[10] The missing rate is calculated as: missing rate = 1 – (N countries with SDG 4 data for group / total N countries within group. Conflict data comes from: https://datacatalog.worldbank.org/search/dataset/003. Dataset accessed: 23 November 2021.

To begin with, the overall missing rate for all SDG 4 targets is around 14 per cent across the two waves (Panel A, column 8). This entails, on average, the lack of data for 2 indicators and 8 outcomes (out of 14 and 58 total, correspondingly). The aggregate missing rate is larger in wave 1 (= 23 per cent) than in wave 2 (= 19 per cent). The missing patterns across SDG 4 targets (columns 1 to 7) is not homogeneous. Highest missing rates are for SDG 4 indicators 4.2 to 4.5 (with rates of at least 35 per cent). Even though the overall missing rate for wave 2 is higher than for wave 1, there is a significant increase in the likelihood of missing data for specific SDG 4 targets in wave 2 compared to wave 1; that is, the rate of missingness for SDG 4.2–4.3 increases by 12–17 per cent for SDG 4.4, and by 10 per cent for SDG 4.6.

Missing rates vary across income groups by SDG targets (Panel B), but not as a whole since missing rates are 10–11 per cent, so there is not a consistent relationship between country income grouping and the degree of missingness. For SDG indicators 4.2–4.3, missing rates are larger for upper-middle- and high-income countries. Yet, for SDG indicators 4.4 (youth and adults attainment rates), 4.5 (percentage of students by language of instruction) and 4.a (school services and school violence-bullying), missing rates are much larger for poorer countries. When it comes to overall regional missing patterns (Panel C), Oceania and, perhaps surprisingly, North America with rates between 21 and 26 per cent, tend to explain the larger part of the overall missing rate. Nevertheless, Africa has missing rates of 50 per cent (SDG 4.4) and 62 per cent (SDG 4.5), and in Asia for SDG 4.2–4.3 missing rates are 36–47 per cent, while in South America the largest missing rates (= 25 per cent) is for SDG 4.5.

Panel D displays missingness by the extent to which the country is conflict affected. Again, as for income, results are mixed, with an overall rate of missingness of 15 per cent for countries falling into the high conflict category, 7 per cent for those in the low conflict category and 14 per cent for non-conflict affected countries. Though, clearly for SDG 4.4, 4.5 and 4.a, countries that have high-conflict status have missingness rates between 1.5 and 2.4 times higher compared to countries non-affected by conflict.

CONCLUSIONS

As we move towards the 2030 deadline, achieving robust monitoring of SDG 4 indicators – and taking action on the data we analyse – requires

Table 2.4 Missing SDG 4 monitoring indicators – missing rates

	4.1	4.2	4.3	4.4	4.5	4.6	4.a	SDG 4–all
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A – time								
All	0.03	0.45	0.50	0.35	0.48	0.27	0.10	0.14
Wave 1	0.04	0.53	0.57	0.38	0.64	0.34	0.29	0.23
Wave 2	0.03	0.65	0.69	0.55	0.55	0.44	0.15	0.19
Panel B – income								
Low income	0.00	0.03	0.07	0.47	0.63	0.03	0.20	0.11
Lower middle	0.00	0.15	0.19	0.44	0.52	0.12	0.10	0.11
Upper middle	0.00	0.31	0.43	0.28	0.46	0.11	0.06	0.10
High	0.06	0.89	0.91	0.20	0.33	0.54	0.06	0.11
Panel C – region								
Africa	0.00	0.12	0.15	0.50	0.62	0.02	0.13	0.10
Asia	0.00	0.36	0.47	0.27	0.38	0.16	0.07	0.10
Europe	0.00	0.78	0.78	0.16	0.14	0.42	0.04	0.08
Oceania	0.11	0.56	0.61	0.42	0.75	0.47	0.11	0.21
North America	0.11	0.79	0.84	0.63	0.84	0.58	0.21	0.26
South America	0.00	0.00	0.08	0.08	0.25	0.00	0.08	0.04
Panel D – conflict								
Conflict affected – high	0.00	0.16	0.18	0.61	0.63	0.16	0.18	0.15
Conflict affected – low	0.00	0.00	0.11	0.22	0.67	0.11	0.11	0.07
Non-conflict affected	0.04	0.54	0.59	0.29	0.43	0.30	0.08	0.14

recognising that missing data in its different forms can be a crucial obstacle. Missing data hamper a robust data picture on what underscores educational inequality. Embedded inequalities are particularly a major bottleneck for resource-constrained education systems, whose countries are furthest away from fulfilling SDG 4. The data for identifying where compounded inequalities are most acute appear to be missing from our current data regime. This chapter demonstrates that (with some additional, but not arduous analysis) it is possible to identify the intersectional drivers of marginalisation. This can offer the chance to turn the identification of missing data into an opportunity to accelerate change towards fulfilling the SDG 4. Thus, missing data assessment can provide a chance

to identify – and address – these groups' marginalisation as a cause for lack of progress on SDG 4.

By relying on a critical assessment of what is included in the SDG 4 monitoring architecture, this chapter has shown how granular marginalisation can be identified through intersecting dimensions of disadvantages operating at different levels which are available from household and learning surveys. The new framework proposed would lead to a better targeting of those groups most disadvantaged and, in doing so, to a more robust monitoring of what the SDG 4 framework attempts to capture – compounded inequalities behind education indicators.

First, neglecting the intersecting disaggregations of SDG 4 indicators according to key drivers of inequality can lead to wider education deprivation. The exclusion of combinations more strongly affects learning indicators than completion/access indicators. For instance, for learning outcomes (target 4.1.1), UIS data neither use two dimension (2D) nor three dimension (3D) layers of learning-driven disadvantages. This omits well-established intersections behind learning such as location-gender-wealth or location-wealth-Indigeneity contributing to the further marginalisation of children at grades 2–3, primary, and in lower secondary. Second, following the heuristic framework proposed, new first layers of inequality become an opportunity to address marginalisation by incorporating future monitoring channels of systemic inequalities based on disability, bullying, preschool attendance, child work, parental education, teacher qualifications and school information (infrastructure, school type, contextual poverty, etc.). Equally, for access/completion indicators, the novel dimensions put forward in this chapter also would allow a more detailed monitoring based on mothers' literacy, region, women's empowerment and community health and development.

REFERENCES

Huisman, J., & Smits, J. (2015). Keeping children in school: Effects of household and context characteristics on school dropout in 363 districts of 30 developing countries. *SAGE Open*, *5*(4), 1–16. doi:10.1177/2158244015609666

OECD. (2018). PISA for development: Technical report. Paris: OECD.

UIS. (2019). SDG4 data digest 2020: How to produce and use the global and thematic education indicators. Paris: UNESCO.

UIS. (2020). SDG4 data digest 2020: Using household survey data to monitor SDG4. Paris: UNESCO.

UNESCO (2020). Global education monitoring report 2020: Inclusion and education: All means all. Paris: UNESCO.

UNESCO (2021). Global education meeting from recovery to accelerating SDG4 progress. Paris: UNESCO.
Willms, J. (2018). Learning divides: Using monitoring data to inform educational policy. Paris: UNESCO.
World Bank. (2019). Ending learning poverty: What will it take? Washington, DC: World Bank.

3. Priorities for missing data and SDG 4: Latin America and the Caribbean

Alejandro Vera, Ernesto Yáñez and Martín Scasso

INTRODUCTION

The current Sustainable Development Goal 4 (SDG 4) monitoring agenda raised expectations of Education Management Information Systems (EMIS) driving a move beyond administrative data, in order to inform policies and plans for long-term educational outcomes (Subosa & West, 2018). The COVID-19 pandemic has also added further complexity to these challenges. However, a lack of data limits governments' and society's capacity to move towards quality education for all. In the case of the Latin America and the Caribbean (LAC) region, missing data takes the form of missing key disadvantaged groups beyond standard disaggregation (e.g., gender, location and socioeconomic status). Typically missed populations from monitoring systems in the region are Indigenous groups, people on the move groups and disabled individuals even if 80 per cent of LAC countries used individual student records in their EMIS.

Even though an EMIS should be comprehensive and systemic by definition (Abdul-Hamid et al., 2017; Subosa & West, 2018), the set of educational data available in a typical country in LAC comes from different sources and is not necessarily integrated into a single information system. In the region, the EMIS's scope tends to be limited as the systems mainly focus on formal education, and they need to be supplemented with administrative records or statistical data from other institutions and sectors (Arias Ortiz et al., 2019).

Focusing on EMIS, the purpose of this chapter is to carry out a general diagnosis of the LAC region's educational data situation, identifying

main data gaps and challenges. The chapter is organized as follows. The next section analyzes the EMIS and other data sources' challenge including all relevant population groups. The third section focuses on the different types and categories of missing educational data in the region. The fourth section analyzes the different education stakeholders' current use of information and the challenges for improvement. The last section summarizes the main challenges identified and provides some recommendations.

MISSING DATA GROUPS: UNDER-REPRESENTED GROUPS OF PEOPLE

The SDG 4 framework has a clear emphasis on promoting inclusive, equitable and quality education, leaving no one behind. Fulfilling this commitment not only requires political will, but also disaggregated data to identify those subgroups left furthest behind. In recent years, LAC countries have progressed towards a timely production of disaggregated and comparable education statistics. However, this is still insufficient to monitor the educational opportunities of some disadvantaged populations. Progress has been mostly concentrated on disaggregation linked to age, gender, geographic location and socioeconomic status. Yet disaggregated statistics about Indigenous peoples and people of African descent, people with disabilities, people on the move,[1] homeless, individuals of diverse sexual orientation and/or other relevant characteristics, has been limited. An in-depth review of the data production processes for these groups has identified common methodological and operational challenges:

- **Identification difficulties**. The absence of standards and the use of inappropriate approaches lead to identification biases.
- **Limited comparability**. The use of different definitions and methodologies not only hinders comparisons at the country level, but also within the country and over time. In many cases, the institutions involved in data collection and production use different definitions, concepts and approaches, which prevent complementarity and comparability.
- **Concentration on access**. Statistics about vulnerable groups tends to focus on access, while data regarding progression, dropout and grad-

[1] It refers to migrants, refugees, asylum seekers and displaced people. See Chapter 9 in this volume for an in-depth discussion of internally displaced people.

uation rates, safety in the educational environment, qualified teachers and/or learning outcomes is notoriously scarce.

- **Minimal data integration/collaboration** among state institutions that generate data on vulnerable groups. As previously mentioned, the lack of data integration is also observed in the Ministries of Education.

A Focus on Students on the Move and Students with Disabilities

Two issues of particular importance for the region have to do with people on the move and those with disabilities. These two groups clearly exemplify the common problems in collecting data on vulnerable individuals, and they have gained relevance in the region's educational agenda.[2] For the first group, the rapid increase in the number of intraregional migrants and the characteristics associated with this migratory flow have resulted in a substantial increase in humanitarian assistance and the emergence of significant challenges for destination countries when it comes to guaranteeing the migrant's right to education.[3] For the second group, there is evidence of very slow progress in the transition to inclusive schools that guarantee the right to education for persons with disabilities (Duryea et al., 2019), despite the fact that most countries in the region are signatories of the United Nations Convention on the Rights of Persons with Disabilities (CRPD).

Recent trends in intraregional mobility have generated tensions and demands of EMIS for new and more disaggregated data. Although in several cases instruments have been modified in EMIS to cover this demand, these changes have been characterized by the use of indirect identification approaches based on nationality or country of birth. However, these changes are of limited use in differentiating between types of migrants and their unique needs. In these cases, administrative data or specific surveys are needed, though they are not generally integrated into EMIS.

---

[2] OREALC/UNESCO Santiago is currently working on the construction of regional monitoring frameworks for these two groups.

[3] Access problems related to the lack of documents, an insufficient educational offer to meet the sudden increase in demand, and difficulties to revalidate studies are some of the issues faced in guaranteeing the right to education for people on the move.

The reliability of the information about students in irregular migration situations is another issue that strains the EMIS. The absence of documents to corroborate aspects such as age or previous studies affects data quality, which in turn affects the estimation of several SDG 4-related indicators such as the parity index for the percentage of over-aged children for the grade (SDG indicator 4.1.5) or the rate of participation in organized learning one year before the official elementary school entry age (SDG indicator 4.2.2). In this context, usual validation mechanisms are inadequate, which has led several Ministries of Education to establish formal and/or ad hoc coordination mechanisms with other agencies, such as migration offices, to cross-check data (UNESCO/OREALC, 2020a).

In the case of people on the move, an efficient educational response requires having high frequency data updates. However, the EMIS collect information on a semi-annual or annual basis, which is insufficient to accurately trace the dynamics of the displaced population and the resulting pressure on the supply and demand of educational services. Even in countries where EMIS are adequately developed, this situation has led – especially in the case of Venezuelan migration – to parallel information gathering that is often not harmonized, coordinated or integrated with the EMIS.

Regarding students with disabilities, most countries collect data from specific censuses, household surveys or by their own EMIS.[4] Despite this, educational data availability is limited. This is reflected, for example, in the low number of countries that showed data related to target 4.5 or indicator 4.a in their voluntary national reviews of progress on the SDG.[5] In the region, data collection on disability is characterized by the use of different approaches for identification and by a marked bias towards "disability" to the detriment of the "social" approach, which is the one recommended by the specialized agencies.[6] This limits comparability and

[4] Only nine countries in the region do not collect data on disabilities, and most of them are Caribbean countries (Yáñez, 2022).

[5] In relation to target 4.5, only one of the 14 reports presents data concerning disability, while for indicator 4.a, only three countries report. The voluntary national reviews considered are from Argentina, Bolivia, Brazil, Chile, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Panama, Paraguay, Peru and Uruguay and were taken from the repository https://sustainabledevelopment.un.org/memberstates (accessed 20 December 2021).

[6] The focus is on identifying whether there is a disability and knowing the type of disability, but not on knowing the degree of difficulty experienced by the student with a disability in his or her social interaction within the school. The

underestimates it. Also, not considering the functional approach, does not enable insights about the true degree of difficulty experienced by people with disabilities in their school participation.

EMIS questions generate biases either due their wording or due to the restrictiveness of the options. For example, the question "Do you have a disability?" means that students with mild or moderate disabilities are not counted. The category "multiple disabilities", without further specification in several countries, implies not knowing anything about students' needs, and reflects a bias for counting rather than for understanding students' needs. In addition, in various countries the information is provided by teachers, who lack adequate training to identify conditions, and their categorization could be subjective, sometimes leading to overestimation or underestimation of disability rates. The weakness of the data translates into a series of difficulties in structuring inclusive and relevant educational practices.

In response to this situation, international cooperation is promoting the use of standardized tools such as the questionnaires developed by the Washington Group and UNICEF (Washington Group, 2020), especially in censuses and surveys in the region. However, less progress has been made in the questionnaires used by EMIS.

EMIS vs Household Surveys to Cover Missing Data Groups

Data about vulnerable groups available in the EMIS depends mostly on the degree of inclusion of these populations in education systems (EMIS do not report on the out-of-school population), while in the case of household surveys it is linked to their design. For example, people with disabilities or on the move are generally under-represented in the surveys, thus implying estimates with significant errors. Also, even in cases where there is a high concentration of migrants in certain areas, it may be difficult to generate a sample that allows extrapolation to the national level. Although there are statistical procedures for correcting this type of error, these require updated sampling frames, which are often unavailable or do not cover these populations. As for EMIS, household surveys also face the challenge of better identifying vulnerable populations. Few surveys incorporate the Expert Group on Refugee and IDP Statistics (EGRIS)

exception is Guatemala, which includes the degree of difficulty that students experience.

recommendations, and migration status is generally based on nationality. Regarding disability identification, it should be recognized that the number of countries integrating modules based on Washington Group developments has increased, although not at the desired pace.[7]

All Spanish-speaking countries (except the Dominican Republic) have included questions about ethnicity in many of their household surveys as well as in the last round of censuses (Global Education Monitoring Report et al., 2020). This progress, nevertheless, has not been equally reflected in the EMIS since data availability on Indigenous students is limited and, in the cases in which information is available, it is of low quality. This makes household censuses and surveys a privileged source of information. However, the region still shows diverse criteria for the operationalization of the Indigenous category (self-identification, mother tongue, language spoken, cultural and/or territorial characteristics), which hinders comparability between countries, and even between different sources within a country.

MISSING DATA TYPES: NEW PRIORITIES AND OLD CHALLENGES

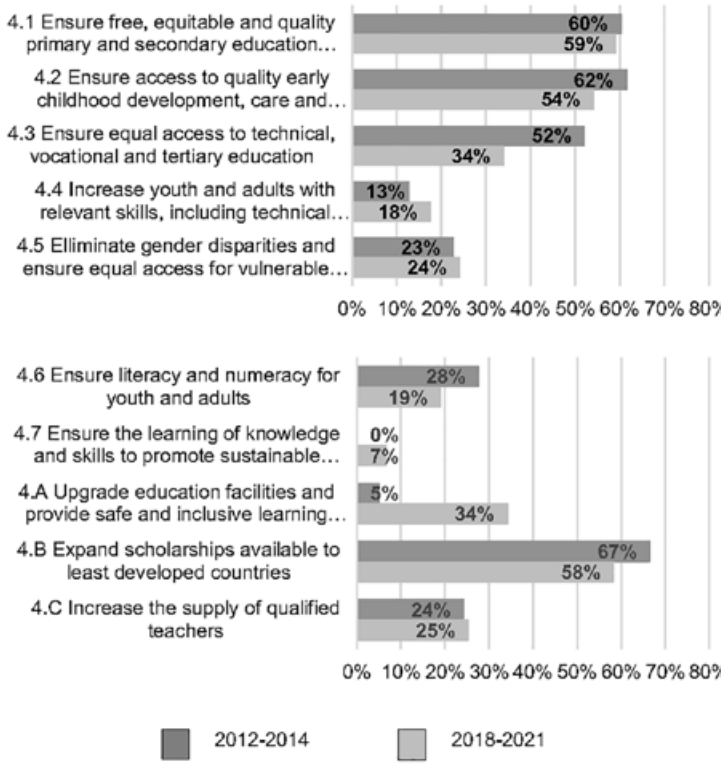
The Regional and National Reporting of Education Data

An assessment of missing data types requires recognizing which aspects of educational planning and decision-making are left out, risking greater exclusion and inequality risks. A first step in this diagnosis can be an analysis of data availability in LAC countries to calculate the global and thematic indicators for SDG 4 monitoring (Figure 3.1). This information reflects the region's monitoring capacity, considering all the available educational data sources (EMIS and others) that countries report to the international agencies responsible for monitoring.[8]

Regional and global SDG 4 monitoring systems show persistent data gaps in LAC for many indicators. The region has greater data production

[7] A review of the 2019 household survey instruments for 18 countries in Latin America shows that only five included questions on disability (Chile, Costa Rica, Honduras, Mexico and Peru).

[8] According to the UNESCO Institute for Statistics (2017), there are four main sources of data for calculating the global and thematic SDG 4 indicators: administrative records, household survey, government expenditure on education and learning outcomes assessment.



Note: This indicator represents the percentage of countries that have information available for each of the "data points" that conform to the global and thematic monitoring framework for each SDG 4 target. A data point is the possible value of each of the 42 indicators in each of the possible disaggregations contemplated in the monitoring framework. Indicator 4.5.1 is excluded because its definition includes all other thematic indicators. For the complete list of indicators: http://tcg.uis.unesco.org/sdg-4-global-and-thematic-indicator-lists/ (accessed 10 December 2021).
Source: Authors' analysis of data from UIS database: https://apiportal.uis.unesco.org/bdds. Accessed 10 December 2021.

Figure 3.1 Percentage of countries with information on each SDG 4 target – 2012–14 and 2018–21

capacity in those dimensions that are typically incorporated into EMIS, but there is a set of fundamental goals that are under-represented, especially in the new dimensions that SDG 4 focuses on. The comparison over

time shows that the region's monitoring capacity has remained stagnant between 2012/14 and 2018/21, and there is even decreasing availability of information to monitor some targets, such as 4.2, 4.3, 4.6 or 4.B. In the 2018/21 period, there is less data for "traditional" indicators, such as the enrolment ratios in pre-primary or tertiary education, youth enrolled in Technical and Vocational Education and Training (TVET), or the literacy rate of youth and adults. This setback is more related to the lack of updated information available at the regional level than to the absence of national data.

When identifying missing data types at a regional level, two issues must be taken into consideration. There is a set of dimensions that are poorly represented because many countries do not produce this information, at least in a way that is adequate for the calculation of indicators. This is typically the case for indicators linked to Early Childhood Educational Development (ECED), learning and skills development for the youth and adult population, or learning to promote sustainable development. Whereas in other dimensions where countries have information, this is not comparable at the regional level. Here, data is not part of the regional monitoring framework but is available for use at the national level. This is typically the case for learning assessments, where relatively more countries have information from national assessment studies, which lack comparability, than those participating in comparable international studies.[9]

At the national level, there are two sets of missing data types: information on certain educational programmes and information on certain dimensions. For the first set, EMIS have traditionally focused on producing data prioritizing educational levels that fall under the umbrella of the Ministries of Education; that is, primary and secondary education, and, in

[9] UNESCO Institute for Statistics indicates that, by 2019, 33 countries in the region regularly applied at least one national assessment focused on primary education, and 27 on lower secondary education (UIS database: https://apiportal.uis.unesco.org/bdds, accessed 10 December 2021). But this information is not comparable at the regional level, so its use is limited to national monitoring systems. Two cross-national learning assessments that produce regular and comparable data in the region are (i) ERCE which focuses on primary education and has a broad coverage in 18 Latin American countries and only one in the Caribbean (https://es.unesco.org/fieldoffice/santiago/llece/ERCE2019, accessed 10 December 2021); and (ii) PISA which produces information on 15-year-old students, in which ten Latin American countries participated in 2018 (https://www.oecd.org/pisa/, accessed December 10, 2021).

most cases, pre-primary (Powell, 2006). The absence of data is associated with educational programmes that depend on different areas of government, which respond to different ministries, and even to different areas of the Ministry of Education itself. Those educational programmes are partially included in the EMIS, and in consequence, some sub-populations attending educational institutions are invisible or under-represented.

On the other hand, national EMIS are weak in covering topics or dimensions that are not part of their traditional scope. When initially developed in the region, these systems focused on recording basic educational resources and processes. The dimensions which were not an education policy priority when EMIS were defined are: school climate and bullying, teacher absenteeism, and language of instruction. For these latter missing data type dimensions, there are other data sources which can help to fill this gap.

A Focus on Early Childhood Education, and Youth and Adult Education

In the LAC region, early childhood education has become an important part of the public and political agenda in recent decades. There is a general consensus on the need to strengthen opportunities for access to this stage of education, and to improve the quality of the education offered in order to boost learning conditions and to increase equity. This is particularly important in the region; LAC has the highest average number of years of compulsory pre-primary education guaranteed in legal frameworks: 1.13 years. In the other regions, the average is less than 0.5 years.[10]

However, most of the countries in the region lack sufficient information to support decision-making to advance towards the achievement of these goals. One of the main reasons for this lack of information is the multisectoral nature of ECED services, which results in the existence of institutions with different interdependencies (e.g., health, education, nutrition, welfare). For example, a recent comparative study of the region identifies seven different types of ECED services: (a) family-centred ECED services, (b) community homes, (c) child development institutions, (d) institutionalized care services in the place of work, (e) institutionalized care services for children without parental care, and (f)

[10] 2020 estimations based on UIS databases (https://apiportal.uis.unesco.org/bdds (accessed 10 December 2021).

early childhood education institutions (Mattioli, 2019).[11] Information is collected in a scattered manner and, for some institutions, there is no system for recording data. In particular, there is a lack of individual student records which could pool different data into a single database; and, sometimes, there is overlapping or double-counting of students.

Another topic that the SDG 4 framework has highlighted is the promotion of broad and flexible learning conditions for the population at different ages and stages of life. Under the concept of "lifelong learning", learning activities are included for people of all ages, in all life contexts and through a variety of modalities. The new frameworks require the inclusion of an array of learning programmes in information systems beyond educational institutions, including formal, informal and non-formal education. These educational and training activities are developed in a multiplicity of locations, depending on different institutions and organizations, many of them outside the traditional EMIS boundaries. Consequently, there is an under-representation of formal educational activities, and there is little or no information on non-formal and informal education (Subosa & West, 2018).

The above examples demonstrate the regional EMIS's limited capacity for change and their constraints in producing information that would allow countries to cover key dimensions of monitoring the right to education, as defined by global and regional frameworks. These obstacles have their origin in the topics and dimensions that were the priority measurement focus when these systems emerged, highlighting the rigidity of the data production processes.

EMIS vs Other Data Sources to Cover Missing Data Types

Standardized learning assessments are complementary data sources that can offset some of the mentioned data gaps. Most of the LAC countries have at least one national assessment focused on primary and/or secondary education, whereas two cross-national learning assessments produce regular and comparable data in the region: ERCE and PISA studies. But there are few experiences in the assessment of competencies among the youth and adult population, or the development of other learning areas,

[11] The website https://siteal.iiep.unesco.org/content/nuevos-perfiles-pais-primera-infancia-siteal contains profiles of 19 Latin American countries and their supply of, regulatory frameworks in, and policies and information available on early childhood education (accessed 1 December 2021).

such as transferable skills, global citizenship education or education for sustainability, all core issues for SDG 4 monitoring. This scenario will improve in the coming years, however. For example, a few countries have recently participated in international studies that assess transferable skills (ICCS, PISA).[12] The ERCE assessment included in its pilot test a module for measuring two soft skills (empathy and self-management) among sixth-grade students. Argentina, Chile, Colombia, Mexico and Peru are developing their own instruments (UNICEF, 2020). Another example for improvement is given by the ERCE assessment, which in 2019 included content related to global citizenship education and education for sustainable development (UNESCO/OREALC, 2020b).

Household surveys are also a very useful resource to compensate for some missing data types.[13] They enable the out-of-school population to be characterized, providing information on participation in educational programmes outside the EMIS boundaries. There are many reasons why household surveys help countries to make progress in the achievement of SDG 4. For example, they measure equity in education, capturing information on non-formal and private education and they provide more accurate estimates of school participation and achievement and literacy rates and educational attainment (UNESCO Institute for Statistics, 2020b). In fact, as of 2017 there were 19 SDG 4 indicators that can be calculated using household surveys (UNESCO Institute for Statistics, 2017). Yet, household surveys are not always able to provide the information needed. For example, although they are the most widely used source of information to estimate the participation and educational attainment rates of the youth/adult population, they do not usually collect data on informal or non-formal programmes.

Another source of information to evaluate is that produced by the Ministries of Education for administrative and management purposes.

[12] ICCS investigates the ways in which young people are preparing to assume their roles as citizens in a world where the contexts of democracy and civic participation continue to change (UNICEF, 2020). PISA includes topics about students' ability to examine issues of local, global and cultural significance; understand and appreciate the perspectives and worldviews of others; engage in open, appropriate and effective interactions across cultures; and take action for collective well-being and sustainable development (OECD, 2020).

[13] In LAC, the main household surveys are the Demographic and Health Survey (DHS) Program from USAID, Multiple Indicator Cluster Survey (MICS) from UNICEF, Living Standards Measurement Study (LSMS), and the national surveys (UNESCO Institute for Statistics, 2020b).

Although these data resources exist and record data that cover educational dimensions other than those included in most EMIS, they are infrequently used for the production of statistics because the adaptation of administrative data into monitoring information is not a straightforward process.

DATA PURPOSES: THE LINK BETWEEN DATA PRODUCTION AND DATA USE

Different Data Sources, Different Purposes

The collection, analysis and dissemination of educational data serve a diverse group of processes, such as management and administration, planning, policy formulation, and monitoring and evaluation. All these processes are interconnected and take place at different locations in education systems, from central to local, and within education institutions themselves (Subosa & West, 2018). Though, as previously mentioned, educational data available in a country comes from different sources. In this sense, the potential uses of educational data will differ depending on the characteristics of each data source and how they have emerged and evolved. Regardless, the boundaries between data sources and purposes are complex. Education stakeholders and government officials tend to combine sources based on their needs and data availability.

The definition and scope of an EMI", as'well as the demands from different data users, evolve over time along with technological development (UNESCO, 2021). These improvements facilitate data collection, enabling more sophisticated processing that opens the door to new uses of educational data. They could also allow EMIS greater decentralization so that data producers, typically actors at the school and lower administrative levels, can benefit from it by immediately accessing and using this information (UNESCO & Global Partnership for Education, 2020).

When it comes to technology, one of the most notable changes in the LAC region's EMIS has been the introduction of individual student-level data, replacing the usual, more aggregated units of measure, such as classrooms or the entire educational institution. This possibility arose thanks to better collection devices, data storage and processing, spreading quickly throughout the region. In 2020, 80 per cent of LAC countries used individual student records in their EMIS, and 58 per cent of countries could retrieve five years of student data based on student ID (UNESCO Institute for Statistics, 2020a).

In some countries, the use of individual records and student IDs also allowed a linkage of EMIS with other databases within and outside of the Ministries of Education. This virtuous process enriched the possibilities in the analysis and use of educational data. A good example of this type of development is the early warning systems to predict and prevent school dropout. These systems combine an array of information (from students' economic and social background, progression and educational outcomes) to identify populations at risk of educational disengagement and dropout, and then to introduce preventive measures at the school level. These systems gained relevance in the COVID-19 pandemic context. Peru and Guatemala, for example, introduced early warning systems as part of their national education response to COVID-19, while Chile, which had this system in place before the pandemic, expanded its coverage (The World Bank, 2021).

Imbalance between Educational Data Production and Use

Due to the improvement of educational information systems, driven by technological developments and the incorporation of individual records, countries currently produce more educational data than ever in history, although this improvement has not been necessarily reflected in the availability of data in the region (as discussed in the previous section). Even if the availability of educational data has brought on new developments and uses of data, especially at the central administration level, there is still room for a more intensive use of data at all management levels and across different stakeholders. In short, an imbalance is observed between the production and use of data.

One reason for this imbalance is the low prioritization of data dissemination within the production cycle by the different agencies in charge of educational data. This translates into delays in dissemination, products that fail to meet the needs of different users, and the lack of training activities in the use and analysis of data for different stakeholders, particularly those who make policy decisions (Abdul-Hamid et al., 2017). This situation, in turn, affects data use as well as the quality and relevance of the data produced by failing to receive good feedback and new demands from users. A virtuous circle of production and data should drive informed decision-making in Ministries of Education and other education agencies (The World Bank, 2021).

The use of data In schools isIer underdeveloped area. In general, the fIow of infoImation in EMIS tends to be unidirectional, from schools to

the centre, with few instances of information being returned to providers: schools, parents and community (UNESCO, 2021). This results in a lower quality of the data and the duplication of processes at the school level, since some of the information needs of the schools themselves could be covered with the data that has already been reported to the EMIS but cannot easily be accessed and processed by school-level administrators.

There are some incipient practices in the region that can offer a step forward in improving educational data use. Within the framework of open data policies in education, LAC countries have implemented several experiences of open data, where the public can access basic educational data and indicators at the school level, including learning outcomes (Brito, 2019). More recently, some countries (e.g., Argentina, Chile and Ecuador) have made educational databases available to the general public and specialized actors in the educational domain, from which innovative data processing can be developed.[14]

KEY MISSING DATA CHALLENGES AND RECOMMENDATIONS FOR THE REGION

In spite of EMIS in LAC countries showing some coverage problems, they continue to be the main source of official educational data. For this reason, the EMIS need to be strengthened and must be at the centre of any strategy to increase the availability and use of educational data, particularly data promoted by international cooperation. Still, it is relevant to recognize that the development of EMIS has different phases and its definition evolves over time. LAC countries are a good example of the diversity in development and complexity of these information systems. Therefore, it is important to align data demands with each country's reality and to adjust expectations accordingly.

This chapter has reviewed the main difficulties affecting educational data in Latin America and the Caribbean, focusing on three broad dimensions: population groups excluded, types of missing data, and limitations in the use of educational data. Based on this diagnosis, below we outline the main regional challenges and offer some recommendations to

[14] See https://www.argentina.gob.ar/educacion/evaluacion-e-informacion-educativa/acceso-especialistas (Argentina), https://dat-osabiertos.mineduc.cl/ (Chile) and https://educacion.gob.ec/datos-abiertos/ (Ecuador). Accessed 20 December 2021.

advance in the objective of having more and better data for educational planningulfilmentlfillment of the SDG 4/Education 2030 Agenda.

Definition of standards and processes to produce new data and to make existing data comparable: In the region there is a marked limitation of data on groups such as people on the move, people with disabilities, and/or Indigenous peoples. A first step to remedy this limitation is to adequately identify these groups. Particularly in the case of persons on the move and/or with disabilities, following the EMIS developments/suggestions made by the Expert Group Refugee and IDP Statistics (EGRIS) and the Washington Group (on disabilities) would not only help to improve identification but also aid comparability, leading to the development of better indicators. Also, because of the importance of ethnicity for educational equality in the region, it is vital to reach a consensus on how to measure it, because few EMIS report this data coherently with other national statistics.

Articulate EMIS with other data sources to broaden the scope of education data: In LAC countries there is a diverse set of data sources that can expand the EMIS scope, covering areas with information gaps. The integration of data sources is essential to compensate for EMIS's limitations. This can be done at different levels of complexity, according to the development of each country's EMIS. For example, a basic level of linkage can be a consistent and harmonized system of indicators from different sources. A more sophisticated integration could be achieved by linking information from different sources at the student level. A very high level of integration may involve an ecosystem of information.

Strengthening and sustainability of regional monitoring systems: The SDG 4/Education 2030 monitoring framework faces a three-layered challenge: to advocate for countries to produce the information they do not have, to promote the adoption of international methodologies for data that is produced in countries but is not comparable, and to compile the data and indicators that countries produce and can be comparable cross-nationally. These monitoring systems need to be reinforced by linking national and regional educational agendas, making the data production sustainable through swift channels of communication and agreement.

The virtuous circle between data use and data quality: One of the greatest weaknesses of the educational information systems in LAC is observed in the final stage of the process, that is, the dissemination of data. Encouraging the use of data is not only crucial because this is the main purpose of having an information system, but it is also a desirable

feature of the data system to improve data quality, through feedback and user demand. In this way, a culture of data use can be generated in the education sector that allows more evidence-informed decision-making.

The SDG 4/Education 2030 Agenda proposes an ambitious set of education goals that calls on countries to promote inclusive and quality educational policies. Monitoring its implementation and the fulfilment of its objectives is a central element of this global agenda and can guide evidence-informed policy decision-making. That requires countries to improve the availability and quality of educational data, to diagnose the different dimensions of the right to education from a holistic perspective, to represent all groups so that no one is left behind, and to make relevant information available to inform better policy decisions.

ACKNOWLEDGEMENTS

The authors appreciate the valuable comments of Daniel Shephard and Juan Cruz Perusia.

REFERENCES

Abdul-Hamid, H., Saraogi, N., & Mintz, S. (2017). *Lessons learned from World Bank education management information system operations: Portfolio review, 1998–2014* (SABER Case Study). World Bank Group.

Arias Ortiz, E., Eusebio, J., Pérez Alfaro, M., Vásquez, M., & Zoido, P. (2019). *From paper to the cloud. Guiding the digital transformation of Education Management and Information Systems (SIGEDs)* (IDB Technical Note 1660). Inter-American Development Bank.

Brito, A. (2019). *Información y transparencia: Cuadros de indicadores de las escuelas en América Latina*. Instituto Internacional de Planeamiento de la Educación de la UNESCO.

Duryea, S., Salazar Salamanca, J., & Pinzón Caicedo, M. (2019). *Somos todos: Inclusión de las personas con discapacidad en América Latina y el Caribe*. Inter-American Development Bank.

Global Education Monitoring Report, Laboratory of Education Research and Innovation for Latin America and the Caribbean & UNESCO/OREALC. (2020). *Global education monitoring report, 2020, Latin America and the Caribbean: Inclusion and education: all means all*. UNESCO.

Mattioli, M. (2019). *Los servicios de atención y educación de la primera infancia en América Latina*. IIPE UNESCO Oficina para América Latina.

OECD. (2020). PISA 2018 results (Volume VI): Are students ready to thrive in an interconnected world? (PISA). OECD.

Powell, M. (2006). Rethinking Education Management Information Systems: Lessons from and options for less developed countries. *InfoDev's working papers series*, *6*.

Subosa, M., & West, M. (2018). *Re-orienting Education Management Information Systems (EMIS) towards inclusive and equitable quality education and life-long learning*. UNESCO Institute for Statistics.

The World Bank. (2021). Acting now to protect the human capital of our children (Report). World Bank.

UNESCO. (2021). Re-imagining the future of Education Management Information Systems "Beyond head counts: Leveraging data systems to support inclusive and effective learning for all". UNESCO.

UNESCO & Global Partnership for Education. (2020). The role of education management information systems in supporting progress towards SDG 4: Recent trends and international experiences. UNESCO.

UNESCO Institute for Statistics. (2017). The quality factor: Strengthening national data to monitor Sustainable Development Goal 4; SDG 4 data digest 2017. UNESCO Institute for Statistics.

UNESCO Institute for Statistics. (2020a). *Data innovation for producing SDG 4 indicators: A global analytical report* (UIS Information Paper, 65 [39], Issue 65). UNESCO Institute for Statistics.

UNESCO Institute for Statistics. (2020b). *SDG 4 data digest 2020: Using household survey data to monitor SDG 4* (SDG 4 Data Digest). UNESCO Institute for Statistics.

UNESCO/OREALC. (2020a). *Diagnóstico rápido sobre la disponibilidad de información sobre estudiantes en situación de movilidad en los sistemas de información educativa*. OREALC/UNESCO Santiago.

UNESCO/OREALC. (2020b). *Qué se espera que aprendan los estudiantes de América Latina y el Caribe? Análisis curricular del Estudio Regional Comparativo y Explicativo (ERCE 2019)*. OREALC/UNESCO Santiago.

UNICEF. (2020). *Importancia del desarrollo de habilidades transferibles en América Latina y el Caribe. Documento de discusión*. Fondo de las Naciones Unidas para la Infancia – Oficina Regional para América Latina y el Caribe.

Washington Group (2020). *An Introduction to the Washington Group on Disability Statistics Question Sets*.

Yáñez, E. (2022). *Disponibilidad de información sobre estudiantes con discapacidad en países de América Latina: sistematización de resultados de encuesta a nueve países de la región. Documento encargado por la Oficina Regional de Educación para América Latina y el Caribe*. OREALC/UNESCO Santiago.

4. Priorities for missing data and SDG 4 in the Asia region

James Shoobridge

INTRODUCTION

Achieving Sustainable Development Goal 4 for education (SDG 4) requires a concerted effort by countries, regions and global organizations to track current progress and address challenges in collecting and producing high quality data (UIS, 2021a). Timely, accurate and complete data is essential to measure learning and identify those who are being left behind, so that policymakers can focus efforts on reaching the most vulnerable and disadvantaged children. These challenges are evident in Asia which comprises almost 60 per cent of the global population and has experienced rapid growth in recent decades. The rapid economic and social development has resulted in many groups of people being left behind (ADB, 2017) which has been exacerbated by recent events such as the COVID-19 pandemic and regional conflicts, with many more children now falling behind in learning even basic skills. This chapter identifies and gives examples of priority areas concerning missing data groups, data types and data purposes in relation to SDG 4 for further investigation in Asia.

This chapter is organized around the following questions:

1. Missing data groups and types: what groups of people and indicators are most systematically under-represented in education data in Asia?
2. Missing data purposes: what purposes do the current educational data system at the national and regional level serve? What is missing?
3. How do missing data issues potentially relate to national and local systems?
4. What are the most dominant challenges faced when linking national, regional and global educational data systems in Asia?

Multiple cross-cutting issues that have been identified by UNESCO (2019) are a relevant starting point to discuss challenges regarding the education data system in Asia. First, data systems only have high coverage of formal, basic education systems and there is a lack of resources (technical, financial and human) that contribute to increasing indicator availability. Second, some metrics related to SDG 4 require new forms of measurement of more complex outcomes than have traditionally been collected by education systems – this often requires linking multiple data sources. Third, there is a lack of data disaggregation that would enable policymaking around issues of educational equity. UNESCO reported in 2016 (UNESCO-UIS, 2016) that only 3 per cent of indicators can be disaggregated by the five proposed characteristics, thus hampering efforts to identify disparities in access, participation and achievement in education.[1] Fourth, there is a need for standardized and comparable methodologies in order to allow for the comparison of learning data both within and between countries. Finally, there is a lack of culture of data sharing and cooperation within and across countries. Within each country, data is collected between different departments and ministries and not generally shared or made easily accessible, which hampers the capacity for countries to report accurately and effectively. Integration with other data sources may warrant better coordination between ministries of education, national statistics offices, other ministries and non-state actors such as civil society organizations (CSOs). All these challenges are relevant to Asia because its size and diversity result in its myriad of data challenges reflecting global challenges.

There has been substantial analytical work undertaken concerning Asian countries' capacity to report on SDG 4. At the regional level there are notable efforts to undertake analysis of country capacity to report against SDG 4. UNESCO undertook an SDG 4 regional readiness assessment on 39 countries in Asia and the South Pacific to help establish a baseline of capacity to report on each SDG 4 indicator (UNESCO-UIS, 2016). The Economic and Social Commission for Asia and the Pacific (ESCAP)[2] produced an annual report on the Asia and Pacific Region's

[1] A more recent report by United Nations Economic and Social Commission for Asia and the Pacific (ESCAP, 2021b) notes that of 12 case study countries in the Asia and Pacific region, none had a complete set of education indicators disaggregated by gender.

[2] ESCAP serves as the United Nations' regional hub promoting cooperation among countries to achieve inclusive and sustainable development and is

progress towards SDG indicators. The annual report provides a high-level comparative regional analysis of SGD 4's 11 core indicators and notes capacity issues. UNESCO, in conjunction with partners, has assessed the ongoing capacity of countries to report on SDG 4 (UNESCO. 2019). Numerous thematic studies have been conducted to evaluate countries' capacity regionally to report against specific SDG 4 indicators such as those relating to learning assessments (ACER, 2019; UNICEF, 2020a).

Comparative regional studies on the capacity of data systems for the reporting of SDG 4 indicators have been carried out. For example, UNICEF undertook a regional study of the use of individual child tracking (UNICEF, 2020b). Other studies have highlighted gaps in the capacity of countries to report comprehensively on child participation in education. UNICEF has undertaken a global study on out-of-school-children between 2012 and 2018 of which the Asia Pacific region was one focal region.[3] A component of the study focused on the systemic capacity of governments to identify and report upon out-of school-children and those at risk of dropping out. The report highlighted the significant gaps in the Asia Pacific region apparent in government data systems for out-of-school children and the children most at risk of dropping out (UNICEF, 2019a). At the country level, many Asian countries, with the assistance of development partners, have undertaken substantial work to analyse and develop their capacity to report on SDG 4 and improve national data systems (BBoS, 2019; GoC-MoEYS, 2019; GoM, 2016; GoN-MoEST, 2019; MoEYS, 2019; UNESCO, 2017b).

The annual regional ESCAP reports provides information on the types of indicators which remain challenging for countries in the Asia Pacific region to report on. The Regional Asia Pacific Education Monitoring Frameworks and SDG 4 report (UIS, 2021a) highlights regional disparities and missing data groups. Thematic reports, such as those concerning learning outcomes and country-specific reports, provide additional information and often highlight particular groups of people who may be under-represented in official national statistics. Understanding the disparities in access, participation, quality and learning outcome is critical to being able to address the challenges facing excluded and disadvantaged

---

the largest regional intergovernmental platform with 53 Member States and nine Associate Members.

[3] The study included Cambodia, China, Indonesia, Lao PDR, Malaysia, Mongolia, Myanmar, the Pacific, Papua New Guinea, Philippines, Thailand, Timor-Leste and Viet Nam.

groups. Therefore, it is necessary to understand which groups of people are under-represented or invisible and in which categories of data they are under-represented. Some examples that are important in the region are highlighted below.

GROUPS UNDER-REPRESENTED

Invisible or Poorly Visible Children

Studies indicate that there are significant numbers of "invisible" children who are under-reported in SDG 4 indicators throughout Asia. UNICEF (2006) defines invisible children as being those children

> who are hardest to reach include those living in the poorest countries and most deprived communities within countries and those facing discrimination on the basis of gender, ethnicity, disability or belonging to an Indigenous group; children caught up in armed conflict or affected by HIV/AIDS; and children who lack a formal identity, who suffer child protection abuses or who are not treated as children.

Displacement and the undocumented status of children are among the major barriers to securing access to quality education for many young children worldwide. These children are often excluded from official databases because of lack of documentation and other reasons relating to lack of protection and care.[4] This group includes children of undocumented and illegal immigrants and migrant workers and children with sub-standard housing such as homeless children and children of herders.

The issue of children without access to basic education is a major cause of concern in Malaysia and the plight of children of migrant parents in the Sabah region of Malaysia highlights some of these challenges. In the context of Sabah, "invisible children" are categorized as: (1) stateless children; (2) dependent children of foreign workers; (3) children of foreigners staying illegally in Sabah; (4) children born from mixed mar-

[4] Including the lack or loss of formal identification; inadequate state protection for children without parental care; the exploitation of children through trafficking and forced labour; and premature entry of children into adult roles such as marriage, hazardous labour and combat. Children affected by these factors include those not registered at birth, refugees and displaced children, orphans, street children, children in detention, children in early marriages, hazardous labour or combat, and trafficked and indentured children (UNICEF, 2006).

riages; (5) children born from traditional and not legally registered marriages; (6) children born soon after marriage; (7) children borne by single mothers; (8) children from Indigenous groups (Allerton, 2020). Due to their absence, or omission from any official government databases, such as the Education Management Information System (EMIS), these children do not exist in government welfare systems and thus no provision is or could be made for their education, health, safety and welfare. There are an estimated 50,000 children in Sabah being stateless and thus without access to health and education services and thus excluded from official government records and statistics.

In the case of Malaysia, the government is working to address this issue and with the support of UNICEF (2019b) has undertaken a study on refugee, stateless and undocumented communities. In the absence of any concrete or official records on the children, parents and households in these "invisible" enclaves, data was procured from alternative learning centres situated in communities or plantations. The study has helped inform the Malaysian government about the extent of invisible children throughout the Sabah region. The government of Malaysia is in the process of taking concrete steps to ensure invisible children are identified and included in the national EMIS and that institutions practising alternative forms of education are also monitored through EMIS.

Children Attending Schools Located in Conflict Regions

Children attending schools in regions of conflict are often excluded from national data systems and are poorly documented. Throughout Asia, regions of conflict often have ethnic minorities which may be seeking some level of autonomy from central governments or may wish to administer their own education systems and deliver specialized curriculum. Examples of such regions include Mindanao in the Philippines (Floresta, 2022), the Xinjiang Uyghur Autonomous Region in Northwest China (Hopper, 2009), and the ethnic regions in Myanmar such as Rakhine state (Myo et al., 2021). In such regions, non-state actors often play a strong role in the delivery of education services. In such cases, non-state authorities may be reluctant to share data with the government as ethnic minorities may be persecuted (World Bank, 2018). Such areas may also present hazards for data collectors and development partners to operate in and thus may also be excluded from government census.

The plight of ethnic children in Myanmar highlights the challenges encountered by children attending schools located in conflict regions.

Conflict affects several states in Myanmar, including Kachin, Northern Shan and Rakhine in what are generally called ethnic areas and the zone of conflicts has grown since 2021 (Council of Foreign Relations, 2022). There are eight types of categories of ethnic education in Myanmar of which only three, ethnic-input schools, mixed schools and hybrid schools are recorded in the government's EMIS (Lall and South, 2013). Schools fully administrated by ethnic groups are poorly represented in government data systems, particularly for schools which are fully resourced by local ethnic groups. Data on enrolment in fully administered ethnic schools is absent from government systems and government administrative systems do not distinguish between regular, ethnic input, mixed and hybrid schools so statistics are not capable of reporting the numbers of schools in each category. Populations in conflict areas in Myanmar are poorly mapped.

The 2014 population census by the World Bank (2018) noted that populations in three conflict areas were not enumerated (MoLIP, 2017).[5] Published figures also do not address the issue of ethnicity or how ethnicity was recorded, as participants did not have the right to self-determine their ethnic group. The census thematic education report also fails to mention ethnic education providers, as school attendance appears to only include government schools and religious institutions. Thus, it is unclear whether those who are out of school are attending ethnic education providers' schools, or whether those who are included as attending school were actually in non-formal ethnic education programmes. The report specifies attendance at "accredited educational institutions", but it does not specify whether that would include ethnic education providers.

There are also issues concerning clear identification of ethnic categories and groups. The official 135 "national races" (*taingyintha*) recognized by the government are considered problematic, representing arbitrary and often imposed categories of identity (MIMU, 2014; Mon

[5] This included an estimate of 1,090,000 persons residing in Rakhine State, 69,800 persons living in Kayin State and 46,600 persons living in Kachin State (MoLIP, 2017, p. 15). In Rakhine State an estimated 34 per cent of the population were not enumerated as members of some communities were not counted because they were not allowed to self-identify using a name that was not recognized by the government. The government made the decision in the interest of security and to avoid the possibility of violence occurring due to intercommunal tension (p. 16).

and Phyo, 2016). The Ministry of Education (MoE) only identifies eight categories of ethnicity on their EMIS forms.

One solution is to monitor the provision of education through the organizations responsible for funding or administering education to these groups. However, the landscape of education providers is complex and variable throughout Myanmar and groups are reluctant to share information with each other. The government, with the aid of the World Bank, was attempting to establish an Ethnic Education Framework Agreement (World Bank, 2020). The framework would allow for non-state actors such as non-governmental organizations (NGOs), CSOs and employers[6] to become accredited Alterative Education Providers (AEP) and to establish an Alternative Education (AE) management database. However, recent events have halted the process and it now looks unlikely that agreement will be reached in the near future.

Children Attending Non-Government Schools

Children attending non-government schools are poorly recorded in many government systems throughout Asia. One main reason for this is that governments have a greater incentive to obtain detailed information on institutions which are publicly funded. Governments are therefore incentivized to collect detailed information on public schools and less incentivized to obtain information on non-government schools. In some cases, non-government schools may be reluctant to submit data to governments for fear of government interference or for financial reasons. For example, in Malaysia the government has deployed Individual Student tracking for government schools but not for non-government schools (UNICEF, 2020b). In Cambodia the MoE does collect public and private education statistics through the same department and figures are published separately leading to disparities in calculations (MoEYS, 2021).[7]

India is an example where students enrolled in private schools may be poorly represented in government systems. India has three broad classi-

---

[6] It is envisioned that various types of organizations and AE providers will be interested in partnering with the MoE, including organizations which operate in Ethnic Armed Organisations (EAO) controlled areas. However, AE providers in this context are not to be confused with Ethnic Basic Education Providers (EBEPs), who primarily represent EAOs and parallel (ethnic) education systems.

[7] This can affect SDG indicators 4.1.3, 4.1.4, 4.1.5, 4.1.6, 4.2.1, 4.2.2, 4.2.3, 4.2.4, 4.a.1, 4.a.3, 4.c.1, 4.c.2, 4.c.3, 4.c.4, 4.c.5, 4.c.6, 4.c.7.

fications of school being public, private aided and private unaided. Data on public schools is collected annually through the national data system called the District Information System on Education (DISE). However, private aided systems are often excluded from the national and state data systems and private unaided schools are very poorly represented in the national database (IZA, 2017).

While the Right to Education Act (RTE) (2009) mandated that no school can run without obtaining a certificate of government recognition, thousands of such schools nevertheless continue to function. District education authorities routinely give warning notices to unrecognized schools each year, threatening to close them down, which suggests they are well aware of many unrecognized schools. The failure to effectively capture data on unaided schools results in an under-estimation in the true size of the private school sector and its impact on learning throughout India (IZA, 2017).

CATEGORIES OF INDICATORS UNDER-REPRESENTED

The readiness assessment conducted by UNESCO in 2016 provides the most comprehensive baseline of missing data throughout Asia and the Pacific Region (UNESCO, 2016). The readiness assessment notes that on average, Asia-Pacific countries were able to collect data for 51 per cent of the 43 indicators and data was available for 44 per cent of the 11 global indicators and 60 per cent of the thematic indicators (UIS, 2017). At the time of the survey, not one country in the Asia-Pacific region reported having the capacity to collect the data to measure all 43 indicators (UNESCO, 2016). The readiness assessment noted that in general, countries reported that it was easier to collect data for indicators based on administrative sources than from other sources, such as learning assessments and household surveys.

Disaggregating Data by SDG 4.5.1

SDG Indicator 4.5.1 requires all data concerning people to be disaggregated by 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).

ESCAP (2021), in a study of five countries in Asia and the Pacific, noted that no country had a complete set of the 12 education indicators

with sex disaggregation. Only 55 per cent of the education indicators were conforming and sex-disaggregated, while another 22 per cent were non-conforming and sex-disaggregated. The remaining 23 per cent are missing entirely or lack sex-disaggregation. In the regional SDG report (ESCAP, 2021b) it was noted that evidence is very limited on gender equality in the region as only two out of nine SDG targets could be measured. This may be attributable to the indicators reflecting the structure of national (or local) education systems and national standards for educational achievement and therefore may not conform to international standards (ESCAP, 2021). As noted above, ethnic (or Indigenous) peoples and those in conflict areas are often under-represented. Disability is often poorly recorded or under-reported. This has been attributed to a lack of disability screening tools, an overall lack of harmonization where such tools do exist, and overall poor disability awareness among teachers (USAID, 2018).

Data on Learning Outcomes

Comprehensive and comparable data on student learning outcomes remains challenging throughout Asia (SDG 4.1.2, 4.2.1). Each country applies its own standards towards measuring learning outcomes. Approaches towards standardized assessment vary for each country and are nationally established.

There have been several notable regional efforts to coordinate and standardize learning assessment. The most notable has been the Southeast Asia Primary Learning Metrics (SEA-PLM) programme. SEA-PLM is a regional assessment that aims to set a common approach to assessing the learning outcomes of students at primary Grade 5, as an alternative to international large-scale assessments and national assessments. SEA-PLM has developed a set of assessment and survey instruments specially designed to suit the context of ASEAN and Southeast Asian Ministers of Education Organization (SEAMEO) member countries (SEAMEO, 2019). However, to date SEA-PLM has only been implemented in six countries[8] and only 2019 data is available (SEA-PLM, 2019).The report highlighted some of the challenges including variation in capacity between countries and differing education systems and durations of schooling.

[8] Cambodia, Lao PDR, Malaysia, Myanmar, Philippines, Viet Nam.

Technical and Vocational Education and Training (TVET) and Non-Formal Sectors

TVET and non-formal education (NFE) have increased in many Asia-Pacific countries in recent times. TVET offers courses recognized as official accreditation while NFE and non-formal learning includes various structured learning situations which do not either have the level of curriculum, syllabus, accreditation and certification associated with formal learning but do have more structure than that associated with informal learning. This definition may vary from country to country and in many Asian countries the delineation between the two forms of education is not clearly defined.

Asian countries do not have a uniform approach to classifying and capturing data on these sectors. TVET and NFE are highly fragmented throughout Asia and most countries have a large number of institutions with different structures of ownership and control such as TVET or education ministries, sectoral ministries, local authorities, employer or employee organizations. Agencies often receive funding from different sources. This results in barriers towards proper monitoring (AFD, 2019).[9]

The example of Cambodia highlights challenges in obtaining reliable data on TVET and NFE apparent in many Asian countries. In Cambodia the Ministry of Labour and Vocational Training (MoLVT) manages some public TVET institutions but does not monitor private institutions. Other ministries, such as the Ministry of National Defense, are responsible for managing specialized TVET institutions. The Ministry of Education is responsible for NFE. There are no mechanisms to support coordination between responsible agencies. Cambodia therefore has to rely on household surveys to obtain data on NFE and TVET which constrains timeliness, frequency and reporting capacity (MoLVT, 2019).

LOCAL SYSTEM CAPACITY TO REPORT ON SDG 4

UNESCO's DQAF reports and other assessments of local capacity, such as assessments of administrative and learning assessment systems, indicate that capacity to report on SDG 4 varies widely throughout Asia.

[9] This encompasses SDG indicators: 4.3.1, 4.3.3, 4.4.1, 4.4.2, 4.4.3, 4.6.1, 4.6.2, 4.6.3 4.b.1, 4.b.2

The variations often relate to specific sectors or capacity to effectively disaggregate data to highlight disparities in education participation, quality and attainment.[10] Gaps in data lead to a failure to adequately plan and provision education towards ensuring an equitable education system in which no child is left behind. Therefore, it is important to understand and address gaps in the data which inhibit equity analysis highlighting disparities.

Some countries such as Cambodia, Timor-Leste and Myanmar require substantial capacity development to report on some SDG 4 indicators and meet requirements for disaggregation of data. Other countries such as Mongolia and Malaysia have significant capacity to report data through all sub-sectors (UNICEF, 2020b).

Accurate costs are difficult to estimate owing to changes in technology and the nature of data systems. UNESCO (2017c, p. 71) place the cost of reporting against SDGs globally as between 2.153 and 2.479 billion USD over a ten-year period. Given the socio-economic diversity throughout Asia, the result is that some countries are better prepared to report against SDG 4 than others.

One significant development throughout Asia has been a move towards national systems which track individual data. Unique national identifiers are assigned to each individual enabling the tacking of children in and out of school through different sector databases such as Health, Child Protection and Education (UNICEF, 2020b). Individual child databases may fail to include some groups of children, such as invisible children, but as children are gradually identified, a more comprehensive and current picture of education within a country emerges. Child tracking databases are currently being deployed in more than 12 countries throughout Asia (UNICEF, 2020b).[11]

There are no direct linkages between data sources regionally. Each country maintains its data systems, some of which are considered to

[10] In accordance with SDG Indicator 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 becomes available) for all education indicators on this list that can be disaggregated.

[11] Including: China, Laos (pilot), Cambodia (pilot), selected states of India such as Andhra Pradesh, Malaysia, Singapore, Thailand, Maldives, Philippines, Mongolia, Taiwan, South Korea.

contain sensitive data. However, there are means by which countries in Asia report data, thus making it comparable. These include:

- **Regional education forums at which participants share data**. Examples include: UNESCO's Asia-Pacific Forum on the Progress Update of the International Commission on the Futures of Education; UNESCO's Asia-Pacific Regional Policy Forum on Early Childhood Care and Education;[12] Association of Southeast Asian Nations (ASEAN) Forum for Vocational Education Management.
- **Regional forums at which participant countries develop capacity to properly report uniformly against SDG 4 indicators**. Partners engage in capacity building workshops regionally such as the United Nations Economic and Social Commission for Asia, UNESCO's East Asia Workshop on Monitoring SDG 4: Education Statistics and Tools, and the Pacific (ESCAP) South-East Asian Forum on the implementation of Sustainable Development Goals.
- **UIS and partner efforts to standardize data gathering and regional reporting on SDG 4 indicators and publish data nationally**. UIS manually distributes Excel spreadsheets annually in a standardized format to capture data relating to SDG 4 indicators. The resulting data is annually published on the UIS website[13] and also made available on other websites such as the World Bank's Open Education Data Bank (called EdStats).[14] Standardized tools have also been developed such as UNESCO's OpenEMIS data initiative (UNESCO, 2018) for collecting and analysing administrative education data, and UIS PacSIMs for collecting and analysing learning assessment data.
- **Developing and applying uniform standards in areas such as Assessment**. The efforts to set a common approach to learning assessment through the SEA-PLM initiative is one example of this (SEAMEO, 2019).

[12] https://bangkok.unesco.org/content/asia-pacific-regional-policy-forum-early-childhood-care-and-education. Accessed 4 September 2023.
[13] http://data.uis.unesco.org/. Accessed 4 September 2023.
[14] https://data.worldbank.org/topic/4. Accessed 4 September 2023.

Making available standardized and SDG 4-compliant OpenSource EMIS solutions. Examples of this include Pacific EMIS[15] and UNESCO's OpenEMIS.[16]

These efforts help to ensure that data is properly comparable. There are notable efforts to standardize some elements of data collection between countries. Dominant challenges include: (i) country sensitivities towards education data comparisons; (ii) sensitive data such as relating to marginalized groups; (iii) diverse systems and methodologies for data collection; (iv) low capacity to apply standards. Standardizing data and ensuring it is comparable will help identify groups within countries being left behind.

CONCLUSION

The Asia region is vast, comprising almost 60 per cent of the global population in 48 countries. SDG 4 places strong emphasis on data generation and monitoring for all Asian countries. The opportunities and challenges in reporting against SDG 4 indicators experienced by Asian countries vary significantly owing to the politically and economically diverse countries within the region. In recent decades, Asia has seen unprecedented economic development and social change. This has resulted in challenges in providing timely and reliable data to report against SDG 4 indicators. These challenges include missing or under-represented groups of people from national datasets and categories of indicators which are challenging to report against.

Groups of people poorly captured in information systems throughout Asia include:

- **Invisible or poorly visible children**. This may include undocumented or displaced children, children of illegal immigrants or migrant workers, or those living in sub-standard housing.
- **Children attending schools in conflict regions**. These children are often members of minority ethnic groups which may be reluctant to share data with government officials.
- **Children attending non-government schools**. Non-government schools may be classified as schools which partially receive government funding or do not receive any government funding. Schools

[15] https://www.pacific-emis.org/. Accessed 4 September 2023.
[16] https://www.openemis.org/. Accessed 4 September 2023.

which do not receive government funding often receive less scrutiny than government-funded schools and may be absent from national systems.

Categories of Indicators which are under-reported in Asia include:

- **Capacity to disaggregate by required parity indices (SDG 4.5.1)**. Challenges include disaggregation of data by gender, disability, wealth and ethnicity.
- **Data on Learning outcomes** to ensure comparable assessment between countries (SDG 4.1.2, 4.2.1).
- **Complete data on the TVET and Non-formal sectors**. These sectors are rapidly growing in many Asian countries to meet the need of multiple economic sectors. However, these sectors are also complex, diverse and managed under complex funding and administrative arrangements often involving multiple agencies.

Many Asian countries are undertaking measures at the national level to address these gaps such as development of their EMIS and inclusion of national student tracking data systems. Exciting opportunities to accelerate capacity are also being driven by rapid technological growth throughout Asia, such as deployment of systems over the internet and accessible via mobile device.

Regionally initiatives are being undertaken to increase data sharing and help countries to further develop their capacity to report on SDG 4 indicators. These include:

- Regional education forums at which participants share data.
- Regional forums at which participant countries develop capacity to properly report uniformly against SDG 4 indicators.
- UIS and partner efforts to standardize data gathering and regional reporting on SDG 4 indicators and publish data nationally.
- Developing and applying uniform standards in areas such as Assessment.
- Making available standardized and SDG 4 compliant OpenSource EMIS solutions.
- Focusing on educating managers in data relevance and use.[17]

[17] For example, cooperative partnerships between the International Institute for Educational Planning (IIEP) and local management training institutions

Many Asian countries are accelerating their investment into information data systems and related infrastructure which is likely to realize impressive gains in the years ahead.

Countries may confirm whether the groups and indicators identified in this chapter are challenges and whether there are other groups of people and SDG 4 indicators which are challenging to report on. Countries may review these data gaps in their context and help identify good examples from Asia which demonstrate overcoming barriers to reporting against SDG 4 indicators.

Timely, accurate and complete education data is essential for policymakers, planners and education stakeholders at all levels of government and for non-government actors participating in the delivery of education services. However, as highlighted in this chapter, there are groups of people who are invisible or poorly represented in national datasets. There are categories of data, such as data on learning outcomes, which are incomplete, or which have other issues affecting the extent to which they can be disaggregated and reported to help identify disparities in the education system. Identifying and addressing these gaps will be paramount towards ensuring resources are channelled towards helping the most vulnerable and ultimately towards addressing inequities in Asia and beyond the Asian region. This chapter presents some examples of data gaps and issues, but is far from exhaustive. Continued assessment of the quality and completeness of education data at all levels of government, regionally and globally, is essential if inequities are to be identified and addressed in a timely manner to help ensure that no child is left behind.

REFERENCES

ACER (2019) *The Southeast Asia Primary Learning Metrics Program: Thinking Globally in a Regional Context*, Southeast Asian Ministries of Education Organisation (SEAMEO), United Nations Children Fund (UNICEF), Australian Council for Education Research (ACER). Available at: https://www.seaplm.org/PUBLICATIONS/regional%20results/the%20southeast%20asia%20primary%20learning%20metrics%20program_%20thinking%20globaly.pdf. Accessed 28 August 2021.

ADB (2017) *Asia-Pacific Sustainable Development Goals Outlook*, Asian Development Bank (ADB), Bangkok, Thailand. Available at: https://www

throughout Asia. See: https://www.iiep.unesco.org/en/our-expertise/training-iiep-unesco. Accessed 4 September 2023.

.adb.org/sites/default/files/publication/232871/asia-pacific-sdgoutlook-2017.pdf Accessed 25 August 2021.

AFD (2019) *Financing TVET: A Comparative Analysis in six Asian Countries*, Notes Techniques Technical Reports No. 56, Agence Française de Développement (AFD). Available at: file:///C:/Users/Jim/Downloads/nt-56-funding-tvet-asian-countries-christine-goncalves3.pdf. Accessed 30 August 2021.

Allerton C. (2020) *Invisible Children? Non- Recognition, Humanitarian Blindness and Other Forms of Ignorance in Sabah, Malaysia*, Critique of Anthropology. Available at: https://www.academia.edu/44731318/Invisible_children_Non_recognition_humanitarian_blindness_and_other_forms_of_ignorance_in_Sabah_Malaysia. Aaccessed 17 January 2023.

BBoS (2019) *National Strategy for the Development of Statistics (NSDS)*, Bangladesh Bureau of Statistics (BBoS), Bangladesh. Available at: http://www.paris21.org/sites/default/files/Bangladesh%20NSDS_2013-23_Final%20Cabinet%20approved.pdf. Accessed 27 August 2021.

Council of Foreign Relations (2022) *Myanmar's Troubled History: Coups, Military Rule, and Ethnic Conflict*. Available at: https://www.cfr.org/backgrounder/myanmar-history-coup-military-rule-ethnic-conflict-rohingya. Accessed 4 September 2023.

ESCAP (2021a) *Bridging the Gap: Mapping Gender Data Availability in Asia and the Pacific*, Technical Report, April. Available at: https://data2x.org/wp-content/uploads/2021/04/BTG-Asia-Technical-Report_FINAL.pdf#page=21&zoom=100,0,0. Accessed 22 October 2021.

ESCAP (2021b) *Asia and the Pacific Sustainable Development Goal (SDG) Progress Report*, United Nations Publication. Available at: https://www.unescap.org/kp/2021/asia-and-pacific-sdg-progress-report-2021. Accessed 27 August 2021.

Floresta, J.K. (2022) *Situating the Role Schools Have Played in the Mindanao Conflict*, ISEAS Perspective 2022/66, 21 June. Available at: https://www.iseas.edu.sg/articles-commentaries/iseas-perspective/2022-66-situating-the-role-schools-have-played-in-the-mindanao-conflict-by-jonamari-kristin-floresta/. Accessed 4 September 2023.

GoC-MoEYS (2019) *National Strategic Development Plan (NSDP) 2019–2023, Government of Cambodia Ministry of Education Youth and Sports (2019)*, Phnom Phen. Available at: http://www.nis.gov.kh/index.php/km/2-uncategorised/84-national-strategy-for-the-development-of-statistics. Accessed 27 August 2021.

GoM (2016) *Republic of the Union of Myanmar, National Strategy for Development of Statistics*, Government of Myanmar (GoM). Available at: file:///C:/Users/Jim/Downloads/NSDS_Core%20Strategies_Eng.pdf. Accessed 27 August 2021.

GoN-MoEST (2019) *Sustainable Development Goal 4: Education 2030 Nepal National Framework*, Government of Nepal Ministry of Education, Science and Technology (GoN-MoEST) Kathmandu, Nepal. Available at: https://moe.gov.np/assets/uploads/files/SDG_4_Nepal_National_Framework_(Final_Print_Ready_Copy)_July_2020.pdf. Accessed 27 August 2021.

Hopper, B. and M. Webber (2009) Migration, Modernisation and Ethnic Estrangement: Uyghur migration to Urumqi, Xinjiang Uyghur Autonomous Region, PRC. *Inner Asia*, 11:10.1163/000000009793066460.

IZA (2017) *The Private Schooling Phenomenon in India, A Review*, Discussion Paper Series, IZA DP No. 10612, March. Available at: https://ftp.iza.org/dp10612.pdf. Accessed 20 October 2021.

Lall, Marie and Ashley South (2013) Comparing Models of Non-State Ethnic Education in Myanmar: The Mon and Karen National Education Regimes. *Journal of Contemporary Asia*, 36 (2), August, 298–321.

MoEYS (2019) Cambodia's Education 2030 Roadmap, Sustainable Development Goal 4, Cambodia Ministry of Education, Youth and Sports (MoEYS). Available at: https://www.moeys.gov.kh/index.php/en/planning/2901.html. Accessed 27 August 2021.

MoEYS (2021) Cambodian Education Statistics, Cambodian Ministry of Education, Youth and Sports (MoEYS). Available at: https://www.moeys.gov.kh/index.php/en/stastic-and-indicator/about-statistic-and-indicator.html. Accessed 30 August 2021.

MoLIP (2017) *Thematic Report on Education Census Report Volume 4-H*, Department of Population, Ministry of Labour, Immigration and Population (MoLIP), and UNFPA. Available at: https://myanmar.unfpa.org/sites/default/files/pub-pdf/4H_Education_0.pdf. Accessed 30 August 2021.

MoLVT (2019) A Review of Cambodia's Technical and Vocational Education and Training (TVET) Data Systems, Asian Development Bank (ADB), not published.

Mon, Y. and P.T. Phyo (2016) After Long Delay, Religious Census Data Proves Less "Sensitive" Than Anticipated. *Myanmar Times*, 22 July. Available at: https://www.iseas.edu.sg/articles-commentaries/iseas-perspective/2022-37-migration-in-post-coup-myanmar-a-critical-determinant-in-shaping-the-countrys-future-by-aung-tun/. Accessed 4 September 2023.

Myanmar Information Management Unit (MIMU) (2014). *Myanmar Census Update September 2014*, Myanmar Information Management Unit.

Myo W., H.N. Htwe, and B. Sae (2021) The Right to Education in a Conflict Affected Area: A Case Study of Access to Basic Education in Rakhine Region, Myanmar. *Journal of Human Rights and Peace Studies*, 7, Supplemental Issue, 133–55.

SEA-PLM (2019) SEA-PLM 2019 Main Regional Report Children's learning in 6 Southeast Asian Countries, UNICEF, Southeast Asia Primary Learning Metrics.

SEAMEO (2019) The Southeast Asia Primary Learning Metrics Program: Thinking Globally in a Regional Context, South East Asian Ministers of Education Organisation (SEAMEO), UNICEF, Australian Council for Educational Research.

UIS (2017) More Than One-Half of Children and Adolescents Are Not Learning Worldwide, UIS fact sheet. Available at: http://uis.unesco.org/sites/default/files/documents/fs46-more-than-half-children-not-learning-en-2017.pdf. Accessed 28 August 2021.

UIS (2021a) *Continental Overview: Bridging Asia-Pacific Education Monitoring Frameworks and SDG 4*, UNESCO Institute of Statistics (UIS), Southeast Asian Ministers of Education Organization (SEAMEO), Association of Southeast Asian Nations and Southeast Asian Ministers of Education Organization (ASEAN-), South Asian Association for Regional Cooperation and Association of Southeast Asian Nations (SAARC), Pacific Community.

UIS (2021b) *Guidelines for Data Collection to Measure SDG 4.7.4 and 4.7.5*, UIS, Montreal. Available at: http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2021/06/Guidelines-for-Data-Collection-to-Measure-SDG-4.7.4-and-4.7.5.pdf. Accessed 22 October 2021.

UNESCO-UIS (2016) *Country Readiness to Monitor SDG 4 Education Targets: Regional Survey for the Asia and Pacific Region*, Information Paper No. 31, United Nations Educational, Scientific and Cultural Organization, Institute for Statistics (UNESCO-UIS), Montreal, Quebec, Ref: UIS/2016/SDG/TD/13. Available at: http://dx.doi.org/10.15220/978-92-9189-202-0-en© UNESCO-UIS 2016. Accessed 25 August 2021.

UNESCO (2017a) *The Quality Factor: Strengthening National Data to Monitor Sustainable Development Goal 4*. SDG 4 Data Digest 2017. Available at: http://uis.unesco.org/sites/default/files/documents/quality-factor-strengthening-national-data-2017-en.pdf. Accessed 26 August 2021.

UNESCO (2017b) Ed-Data Quality Assessment Framework (Ed-DQAF) to Evaluate Administrative Routine Data Systems, Manual for the Conduct of an Evaluation by a National Technical Team. Available at: http://uis.unesco.org/sites/default/files/documents/training-workshop-manual-data-quality-assessment-framework-2017-en_0.pdf. Accessed 27 August 2021.

UNESCO (2017c) *The Data Revolution in Education*, UNESCO, Information Paper No. 39, March.

UNESCO (2018) OpenEMIS Concept Note – Equity in Education Tracking and Monitoring Equity in Education Data Using an Open Education Management Information System (OpenEMIS), UNESCO, Paris. Available at: https://www.openemis.org/wp-content/uploads/2018/04/OpenEMIS_Concept_Note_Gender_en.docx. Accessed 30 August 2021.

UNESCO (2019) *Beyond Commitments, How Countries Implement SDG 4*, UNESCO, Global Education Monitoring Report. Available at: https://en.unesco.org/gem-report/node/3093. Accessed 28 August 2021.

UNICEF (2006) *The State of the World's Children in 2006. Excluded and Invisible*, United Nations International Children's Fund (UNICEF), New York. Available at: https://www.unicef.org/media/84806/file/SOWC-2006.pdf. Accessed 30 August 2021.

UNICEF (2018) *A Statistical Profile of Child Protection in Cambodia*, United Nations International Children's Fund (UNICEF). Available at: https://www.unicef.org/cambodia/reports/statistical-profile-child-protection-cambodia. Accessed 30 August 2021.

UNICEF (2019a) Learning Against the Odds. Evidence and Policies to Support All Out-of-School Children and Adolescents in East Asia and Pacific, United Nations Girls' Education Initiative (UNGEI), August. Available at: https://

www.unicef.org/eap/media/3816/file/out%20of%20school.pdf. Accessed 30 August 2021.

UNICEF (2019b) Children Out of School, the Sabah Context, UNICEF, Ministry of Education Malaysia. Available at: https://www.unicef.org/malaysia/media/921/file/Out%20of%20School%20children%20%20(OOSCI)%20Accessible%20version.pdf. Accessed 22 October 2021.

UNICEF (2020a) *SEA-PLM 2019 Main Regional Report Children's Learning in 6 Southeast Asian Countries*, United Nations Children Fund (UNICEF), South Asia Primary Learning Metrics (SEA-PLM). Available at: https://www.unicef.org/eap/media/7356/file/SEA-PLM%202019%20Main%20Regional%20Report.pdf. Accessed 28 August 2021.

UNICEF (2020b) *Review of Education Management Information Systems (EMIS) That Track Individual Student Data, Summary Report*, Oxford Policy Management (OPM), United Nations Children Fund (UNICEF). Available at: https://www.unicef.org/eap/media/6031/file/EMIS%20summary.pdf. Accessed 28 August 2021.

USAID (2018) Cambodia Situational Analysis of the Education of Children with Disabilities in Cambodia Report. May, p. 33.

World Bank (2018) Ethnic Education in Myanmar Literature Review, *World Bank Literature Review*, 1.

World Bank (2020) International Development Association, Project Appraisal Document on a Proposed Credit in the Amount of SDR72.4 Million (US$100 Million Equivalent) and Proposed Grants in the Amount of US$70 Million from the European Union to the Republic of the Union of Myanmar for an Inclusive Access and Quality Education Project, World Bank, 5 February.

5. Priorities for missing data on SDG 4 in the Arab Region

Karma El Hassan

INTRODUCTION: ARAB REGION CONTEXT

The Arab States have reconfirmed their commitment to the Sustainable Development Goals (SDGs) through regular regional meetings; however, many of the countries in the region continue to struggle to make progress towards the SDG 4 goals. Numerous challenges face the region, including heterogeneous contexts, weak data systems, and instability affecting multiple countries in the region. This chapter identifies priority gaps within education data systems and their use in the region alongside recommendations for improvement.

The Arab Region is composed of 22 States located in North Africa and Western Asia and the Horn of Africa. It has four subregional groupings, and they are presented in Table 5.1.

Table 5.1 Arab subregions and countries

Subregion	Countries
Gulf Cooperation Council (GCC)	Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates
Mashreq	Egypt, Iraq, Jordan, Lebanon, the State of Palestine, and the Syrian Arab Republic
Maghreb	Algeria, Libya, Morocco, and Tunisia
Arab least developed countries (LDCs)	Comoros, Djibouti, Mauritania, Somalia, the Sudan, and Yemen

Arab Region and SDGs

Since 2015, efforts to embrace the 2030 Agenda and the SDGs are on the increase across the Arab Region. Governments are changing the institutional setup for related policy implementation, aligning national priorities with the SDGs, and launching voluntary national reviews. However, despite all these efforts the region is not on track to achieve the 2030 Agenda. The region lags in many critical targets and has uneven and worrying rates of achievement (Arab Forum for Sustainable Development: ESCWA, 2020) as there is wide disparity in the development performance across countries. The region is beset with several challenges and has had a decade characterized by civil unrest, regime change, protracted conflict, and crises of displacement.

Arab Region and SDG 4

With respect to SDG 4, Arab Member States have shown firm commitment to transforming national education systems in line with the Education 2030 Agenda and have gathered at four successive Arab Regional Meetings on Education 2030 (AR-MED) since 2015: Cairo in December 2015; Dubai in March 2017, Dead Sea, Jordan, in November 2018; and virtually in Beirut, in July 2021 (UNESCO, 2022). Each AR-MED resulted in concrete action points through a Roadmap, that were intended to enable both countries and partners to generate national momentum, unpack and contextualize the Education 2030 Agenda and its ten targets and indicators, and promote multi-sectoral and stakeholder partnerships.

The year 2020 presented an auspicious opportunity as it launched the "Decade of Action: Ten Years to Transform our World", a call for accelerating sustainable solutions to achieve SDGs, including SDG 4 on quality education. Nevertheless, the year 2020 faced new challenges primarily in relation to the COVID-19 global pandemic that forced school closures and stark declines in both domestic and international economies. Most of the countries in the region have relied on alternative learning modalities, primarily using technology, and switching between online/remote and hybrid learning. However, much of the practical training for skills development was interrupted. This had the effect of exposing and deepening existing socioeconomic and other inequalities that lead to the exclusion of the poorest, most disadvantaged, and marginalized, as well as those under fragile situations in the region (Venkatraman, 2021a). The

July 2021 meeting focused on the need for new data, quality standards, and coordination between ministries in addition to monitoring challenges especially in relation to the COVID-19 global pandemic and for integrating disability in monitoring education.

Despite expanding enrolment and increased investment, education is falling short of its transformative potential in the Arab Region. Prior to COVID-19, the region was showing steady progress based on available data from national reports. However, three main priority areas of action remain: displacement, migration, and marginalized vulnerable groups; quality and relevance of education; and education financing (AR-MED III, 2018).

Data collection and disaggregation are limited and there are data for only six of the 11 core SDG 4 indicators (ESCWA, 2020). Only 47 per cent of the countries in the region are collecting data required to produce the 11 global indicators and 59 per cent of the countries are collecting data for all 32 additional thematic indicators (Venkatraman, 2021a). Four countries, Egypt, Jordan, Morocco, and Tunisia, monitor at least 80 per cent of the formerly labelled tier 1 indicators. Meanwhile, in countries experiencing conflict, such as Syria and Yemen, the best efforts yield about 20 per cent of such indicators being monitored. The relative paucity of statistical data is more evident when considering disaggregated data, with 43 per cent of disaggregated indicators simply unavailable. For example, only Egypt and Palestine report on the enrolment of persons with disabilities. Indeed, in the region, only Egypt, Algeria, Qatar, and Morocco use disaggregated datasets, while conflict countries have virtually no such information (Wahbe, 2021).

MISSING DATA TYPES AND GROUPS

Divergent Development Patterns

The difficulty with education indicators is compounded when one considers the vastly divergent development patterns across the Arab States. Although the region is united by language, culture, and history, it also contains great diversity in resources, income levels, political traditions, social outlooks, and cultural practices. This makes comparability within the region difficult, especially due to varying methods used for the calculation of education indicators and the different adopted standards and expectations of each country. Identifying data availability and gaps for all SDG 4 indicators can be challenging due to a lack of national stand-

ard guiding documents on national indicators, data sources, and quality requirements (Venkatraman, 2021b). Data is not available at all for five of the indicators; and for those with data, averages across Arab States are misleading because of the divergent pattern of development in the Arab world (ESCWA, 2020). Taking indicator 4.2.2 as an example, the participation rate in organized learning one year before the official primary entry age (by sex) has an average of 47.2 per cent across Arab countries but ranges from 5.6 per cent to 70.6 per cent for individual countries. The latter (maximum) participation rate is in line with the world average while for many countries, for example, least developed countries (LDCs) such as Djibouti and Yemen, the rate falls to less than 10 per cent.

Data Context in the Arab Region

Data in the Arab Region suffers from serious limitations. Most notably, there is limited availability of systematic, reliable, and valid data subject to disaggregation in order to monitor global and thematic indicators. The ESCWA March 2022 report on data availability confirms that large gaps in data availability still exist in several areas related to sustainable development in Arab countries. Key challenges include the quality and frequency of the statistical information produced nationally, and their international comparability, as well as data transparency and accessibility. The statistical apparatus needed to make a judgement on the achievement of the SDGs is seriously deficient (ESCWA, March 2021). Despite a lack of SDG monitoring data at national levels, there is a growing amount of project-based data grounded in small, non-representative surveys (Luomi, 2020).

The World Bank's data capacity score (DPR), a composite indicator based on national statistical capacity related to methodology, data sources, periodicity, and timeliness, indicates a low capacity of the 16 countries covered in the Middle East and North Africa (MENA) (SDG Index and Dashboard Report: Luomi et al., 2019). In 2020 the DPR averaged 54 out of 100 for MENA and 55.5 for Arab countries with scores ranging from 22 for Syria to 82 for Egypt. Moreover, there is the absence of a cohesive and integrated statistical system in most Arab countries and there is limited use of other sources of data. As no publicly available regional datasets were identified, all new indicators rely on global datasets, many of which have important gaps for the Arab region (Luomi et al., 2019).

In addition to limited collection and availability of data, the nature of SDG 4 indicators poses a challenge. Many indicators are new conceptually, still under study, and/or have a wide scope and multiple disaggregation variables (ESCWA, 2020). As mentioned above, there is a limited level of disaggregated data in the Arab Region, in fact 43 per cent of disaggregated indicators are simply not available, especially in conflict countries where no such information is available. Moreover, the pandemic has highlighted the need for new indicators that measure rapid assessments at any time (ESCWA, 2022).

Underrepresented Indicators

Data availability and quality in the Arab Region is a major issue as described above. Because of the above situation, the present section will highlight four of the most salient gaps in data by indicator to measure SDG 4 in the region: data on the affordability of education, data on holistic educational goals (SDG 4.7), learning achievement data (SDG 4.1.1), and data on early childhood education (SDG 4.2).

There is a lack of data that captures the affordability of education. During the first five years of the SDG agenda, only 68 per cent of countries in the Arab Region reported any data on SDG 4.5.4 on educational expenditures (UIS, 2021). This is critical in the Arab Region where there are large disparities between public and private schooling, and where most families spend substantially on education. An example can be cited from Morocco where there has been a threefold increase in the number of students in private schools in the last 15 years, with most private schools located in urban centres (ESCWA, 2020). In Mauritania, a least developed country, the growth rate of students in private education was double the rate in the public sector between the school years of 2011–12 and 2016–17. Public schools are closing, and their land is being sold as the government is short on funding (ESCWA, 2020). Furthermore, some public school systems continue to charge nominal fees either to all students or to subsets of students (e.g., specific nationalities are charged fees in some Arab States). However, the data on the cost implications for education systems and families is substandard and national indicators for "free and compulsory" schooling per year (4.1.7) and educational expenditure (4.5.4) are insufficient and often hide the continued cost of schooling borne by families and learners.

Current data also fall short of providing a measure of the holistic goals of education in terms of the extent to which education contributes

to building inclusive and equitable societies (SDG 4.7). Only 50 per cent of Arab States reported a single indicator under this target during the first five years of the SDG agenda (UIS, 2021). This is particularly important for SDG 4 with its focus on the relevance of learning for social and civic life. The measurement of SDG 4.7 and achievements under the concepts of Education for Sustainable Development (ESD) and Global Citizenship Education (GCED) have posed several challenges in terms of the development of indicators to monitor progress, both at the global and Arab Region levels. This target has a wide range of concepts and addresses a range of topics for education to cover including ESD, human rights, gender equality, promotion of a culture of peace and non-violence, GCED, and the appreciation of cultural diversity (Venkatraman, 2021a). It presents difficulties in developing a consensual monitoring framework for it and requires an interdisciplinary approach (Venkatraman, 2021a). Standards and methodological tools needed for specific knowledge, skills, values, and behaviours associated with ESD and GCED and the proficiency levels to be achieved are yet to be globally agreed upon. These standards consist of definitions, concepts, classification systems, and methodologies that help improve harmonization and comparability of official statistics. The global indicator for this target is largely missing for Arab countries and some concepts embedded in the 4.7 target remain politically contentious in several countries. Only a few data points are available for the thematic indicator that measure the proficiency in knowledge of environmental science and geoscience and the percentage of schools that provide life skills-based human immunodeficiciency virus (HIV) and sexuality education. Only a few countries have reported that life skills-based HIV and sexuality education are available in schools (UIS, 2021).

Learning achievement (4.1.1) data remains sparce and unstandardized in the region. Only 59 per cent of Arab States have any reported data in the UIS dataset for the first five years of the SDG agenda. Although data from national learning assessments are reported, performance levels from national assessments are not directly comparable because each country sets its own standards based on national curricula and expectations by level of education and subject. Learning assessments are comparable only for those participating in cross-national international assessments

like PISA, TIMSS, and PIRLS.[1] The Arab States also lack a regional assessment, such as those in Africa and Latin America. So, there is a need for more comparable data on learning achievement that is based on standard metadata and in accordance with international recommendations and standards. For example, on indicator 4.1.1, proportion of children at the end of lower secondary achieving at least a minimum proficiency level in mathematics, the least developed countries in the region could not report a score while Mashreq countries reported an average proportions of 0.48, Gulf Coperation Council (GCC) reported 0.41, and the Maghreb reported 0.37, confirming the large subregional differences (ESCWA, 2020).

As for early childhood education, data collected using household surveys for this target is scarce across the Arab States. Only five data points are available for the global targets on early childhood development and attendance, however administrative data is available, allowing us to monitor enrolment rates on early childhood (UIS, 2021). An additional challenge for the region is that participation rates in early childhood education (4.2.2) are not comparable due to varying methods used for calculation, so these also need to be standardized. Such demands on monitoring require creating and maintaining data quality standards and the development of new measurement methodologies. New data sources will have to be identified and the already existing ones will have to be assessed, extended, improved, and integrated with other data sources.

There are other gaps in missing data like those pertaining to indicator 4.a.1 on providing safe environments and the unavailability of disaggregated datasets in some countries, and the lack of data on the participation rate of youth and adults in formal and non-formal education and training, but the above four are of utmost importance as they help attain SDG 4 outcomes related to acquisition of foundational and relevant skills. Arab countries have a firm commitment to transforming their educational systems in line with the Education 2030 Agenda and for this reason gathered in four Arab Regional successive meetings on Education 2030 (AR-MED) and have worked on national reports as previously mentioned. Building the relevant capacities and ensuring the comprehensiveness of data coverage are very important for monitoring SDG 4 progress. This would help address the concern of making it more inclusive to ensure that "no one is left behind" in every sense of the term,

[1] Seven countries in the region participate for Grade 4 and ten countries participate for Grade 8.

such as educational access, equity, equality, and inclusion and these are prime targets for the region (UNESCO, 2022).

Underrepresented Groups

Due to the sociopolitical situation in the region, the region suffers from increasing inequality and shrinking civic space. Amid widespread conflict and crises in the region, many children in Arab countries continue to be at risk of restricted educational access, low quality, early dropout, and unsafe educational environments. The impact cuts across all dimensions of development (ESCWA, 2020). The pandemic deepened existing inequalities in the region, including socioeconomic ones that lead to the exclusion of the poorest, most disadvantaged and marginalized, as well as those living under fragile situations. The region has witnessed a turn away from public schooling and a move towards privatization and commodification and, while a solution for some, this has increased inequality (ESCWA, 2020). Socially responsible statistics should ensure that everyone is counted, so that those who tend to be left behind can have their needs made visible. In this respect, SDG 4 data and the structure of current data system in the Arab Region lack comparable and disaggregated data on people with disabilities, older persons, women, girls, refugees, migrants, displaced persons, and other vulnerable and marginalized groups, and pose a critical development issue (ESCWA, 2022). For example, in 2021 MENA accounts for 21 per cent of all internally displaced persons (IDMC, 2022) and 11 per cent of all refugees (UNHCR, 2022) despite only having 5 per cent of the global population (UNDESA, 2022). Some of the underrepresented groups include the following generations of refugees and internally displaced people: children from low-income families; persons with disabilities; and youth who are not in employment, education, or training.

Generations of refugees and internally displaced people cannot access education for reasons ranging from discrimination, lack of status and official documentation, and difficulty in adapting to unfamiliar curricula and languages. Many refugee and internally displaced children also stay out of school to work and help support their families (ESCWA, 2020). Across the region and already prior to the pandemic, an estimated 15 million children were out of school, mainly due to conflict and protracted crises (UNICEF, 2019). These numbers are often estimated imprecisely as data on such out-of-school, conflict-affected children is largely missing. In Mosul in 2017, 74,000 out of 141,000 school-age children

(50 per cent) could not access any form of education. Meanwhile, by 2018, 489 schools in Libya and around 2,000 schools in Yemen had been damaged or destroyed by conflict (OCHA, 2019). Syria remains the largest displacement situation in the world, while Yemen remains the worst humanitarian crisis globally, with 66 per cent of the population requiring humanitarian assistance. See Chapter 9 in this volume for an in-depth discussion of the issue of missing education data for internally displaced children.

The poorest children in some countries of the region are less likely to attend school and less likely to complete primary school compared to children in higher-income groups. This gap is amplified in rural areas. Despite increased national investment in education, many rural communities lack the infrastructure, physical or educational, to extend quality, inclusive and equitable education. Rural-urban differences become particularly pronounced when it comes to the probability of reaching secondary education. While the probability of reaching secondary school is quite high for the most advantaged girls and boys in most countries, the rate for the least advantaged girls falls to 28 per cent in Yemen, 20 per cent in the Syrian Arab Republic, and 12 per cent in Iraq. In Morocco, for example, where the rural population has less access to and lower quality of education services and institutions, 26 per cent of children drop out in the last grade of primary (Grade 6), compared to 1 per cent in urban areas (ESCWA, 2020). The adjusted location parity index (LPIA) for proficiency level in reading at the end of lower secondary education, which shows the relation between urban and rural children, revealed great disparity and a more favourable situation for urban children. In terms of the out-of-school rate, a few countries present rates higher than 1.03 GPIA (adjusted gender parity index), meaning that girls are more likely to be out of school compared to boys in these countries. In Gulf countries, 100 per cent of primary schools have access to the internet for pedagogical purposes. The results vary for other subregions. In the primary level in Maghreb countries, Morocco reported that 80 per cent of schools are equipped with computers and the internet while only 5.3 per cent of primary schools in Algeria had access to the internet in 2020. Access to basic services (computers and internet) was high for secondary schools (above 85 per cent) in all Maghreb countries in 2020 except for Mauritania where all schools in 2016 had computers but none of them had internet access. In the Mashreq, all four countries reported a high proportion of schools with computers and they also have high proportions of connected schools. Egypt is the outlier in this indicator, reporting that

only 50 per cent of the secondary schools have an internet link available for the students. No data exists for the least developed countries (UIS, 2021).

In many countries of the region, persons with disabilities are particularly excluded from all education cycles, and especially post-primary education. This is even more the case for girls and women with disabilities in rural areas. For example, only 1.8 per cent of girls and young women with disabilities aged 15 to 24 in rural areas of Yemen attend school (ESCWA, 2020). Literacy rates are lower for persons with disabilities across the region. In Oman, for example, 87 per cent of persons without disabilities are literate compared to only 31.2 per cent of persons with disabilities (ESCWA, 2018). The current data structure in the Arab Region, with its inadequate apparatus and lack of disaggregation, does not accommodate the identification and inclusion of learners with disabilities. Only a few data points are available on the proportion of schools that are disability-friendly at the primary level and fewer at higher educational levels. The information available for Gulf countries is that five out of six schools are adapted to children with disabilities. Morocco and Palestine are the only countries in their respective subregions with data on this indicator. In Morocco, around 20 per cent of the faciIities in primary schools are adapted (in 2020). In Palestine, the proportion of adapted schools increased from 34 per cent in 2016 to 54 per cent in 2020 (UIS, 2021). As is the case in other regions (i.e., see Chapter 3 for Latin America and the Caribbean), there is a challenge in producing comparable data on persons with disabilities due to the varying definitions and understanding of the term "disability". This situation has led to obstacles of inconsistent data collection methods that limit disaggregation – even within the same country.

To help in this issue, The Washington Group on Disability Statistics developed a set of questions for use in household surveys and censuses to consistently collect information on persons with a disability (Venkatraman, 2021b). At the same time, the Statistics Division at ESCWA embarked on a project with Arab countries to compile, verify, and harmonize national data to the extent possible to allow for better comparability and improve national reporting. The outcome of this work resulted in several publications and a Regional Guidebook to Improve Disability Data Collection and Analysis in the Arab Countries (ESCWA, 2018). Furthermore, in 2019, the Statistics Division at ESCWA produced the first disability framework: 115 indicators to bridge the gap between policy and statistics. It is the first three-dimensional disability frame-

work that maps disability-related indicators to three major development frameworks. The framework will, through improved disability statistics, provide a clearer method for comparing persons with disabilities and those with no disabilities. It will also assist data producers to provide valuable and better information for policymakers to ensure that all people with disabilities everywhere and anywhere are indeed included. Simultaneously, the first regional database on disability statistics was developed, which will be updated from recent household surveys and censuses implemented by member countries. A dashboard presenting country profiles through a disability lens is also being finalized (ESCWA, 2019).

Young people who are not in employment, education, or training (NEET) are a rising and disquieting phenomenon in many countries in the region accounting for 18 per cent of Arab youth, 10 per cent for males and 27 per cent for females. Around 45 per cent of young people in Yemen, 32 per cent in the State of Palestine, and 28 per cent in Egypt are considered in this category (UNICEF, 2019). Despite this need, the educational indicators for young people and adults continue to be forgotten and missing. For example, during the first five years of the SDG 4 agenda, only 9 per cent of Arab States provided any data on the participation of youth and adults in non-formal education or training (4.3.1), only 18 per cent provided any data on tertiary enrolment (4.3.2), and only 59 per cent provided any data on youth participation in technical and vocational education and training (4.3.3) (UIS, 2021). This risks contributing to an increase in young people who are frustrated, idle, and unable to integrate into formal, sustained, and productive employment.

Disparities and data gaps also coincide when we turn to indicators focusing on if youth and adults have fundamental numeracy/literacy skills (4.6) and relevant skills for accessing decent work (4.4). The data coverage is poor, only 64 per cent of Arab States reported any data on any indicator linked to either of these targets between 2015 and 2020 (UIS, 2021). These data gaps are all the more worrisome given the disparities revealed in the existing data. The proportion of youth and adults with basic ICT skills is higher in Bahrain, Jordan, and the United Arab Emirates (60 per cent to 90 per cent) than in Egypt and Morocco (20 per cent to 50 per cent). Meanwhile, the proportion becomes much lower in Djibouti and the Sudan (less than 20 per cent) (ESCWA, 2019). Gulf countries have the highest proportion of skilled youth and adults, and the figures are consistently above those in the other regions. As an example, 73 per cent of youth and adults in the Gulf region have enough

skills to attach a file to an email, while this proportion is 36 per cent in the Mashreq, the second highest average.

The above groups represent those underrepresented in Arab Regional data on SDG 4, yet the above issues are more accentuated in LDCs in the region as evident from the above data. The LDCs simultaneously have the worst data coverage along these dimensions and report the lowest performance on indicators when data is available. This creates a worrying situation in which those countries that most need data to target scarce resources and reach the most vulnerable are least likely to have such data. For example, the adjusted net enrolment rate (one year before the official primary entry age) was 13.4 per cent for Djibouti in 2020 and 39.9 per cent for Sudan in 2018. In addition, participation in vocational education was 2.6 per cent for Djibouti in 2021 and gross enrolment ratio for tertiary education was 16.9 per cent in 2015 for Sudan. In LDCs, the highest proportion of skilled youth and adults was 11.9 per cent for copying files (a simple task) and 3.05 per cent (the lowest) for complex tasks like knowledge of computer programming languages. Finally, among the three LDCs, the gender parity achieved in 2016 was lost in the last couple of years. For early childhood level, Djibouti was favouring more boys in 2019 than it was in 2015. In 2016, Yemen reported a parity rate of 0.89 for this level. As another worrying example, no data is available for SDG 4.a on the school environment in LDCs.

MISSING PURPOSEFUL AND PROCESS-ORIENTED DATA

The above issues in missing data and groups are aggravated by missing purposes and processes. For example, as of December 2019, only half of the official indicators had an established methodology and data for at least 50 per cent of countries in the region, and almost half of the targets were not quantified, which makes their tracking difficult (ESCWA, 2020). The following paragraphs outline major issues that are present in the processes and procedures of the education systems in the region.

Inadequate policy frameworks, policy gaps, lack of integration and alignment of policies with development and SDG framework and the 2030 Agenda are characterizing features of the education systems in the Arab Region. The region needs an urgent overhaul of policies to address intersectional inequality, environmental degradation, and youth unemployment, among other issues. Missing or inadequate policies derail efforts to address the mammoth challenges facing the region in achieving

the SDGs (ESCWA, 2020). At the same time, there is a need for localizing and disaggregating indicators based on each country's own national contexts and priorities (UNESCO, 2022).

The content that education systems deliver remains far from being transformative as envisioned under SDG 4. Curricula are limited and rarely innovative; the teaching profession is undervalued and underpaid; the humanities and social sciences are not prioritized; and learning continues to be mostly by rote at least in public schools. The emphasis on end-of-year testing devalues problem-solving and analytical skills and emphasizes the ability to reproduce information as delivered by teachers. Thus, the system does not encourage active engagement with knowledge and its production, rather than simply acquiring it and this may result in lower capacity to produce robust local data systems. Moreover, the SDG 4 vision of promoting human rights, gender equality, and sustainability is not well integrated in the curriculum (ESCWA, 2020; UNESCO, 2022).

There is limited investment in research and development, low technological productivity and knowledge production, and lack of frameworks and standards for emerging areas in education that are critical for the future of children and youth (ICT, ESD, GCED, etc.). Only a few countries have ensured the minimum investment in education of 15 per cent of total government expenditures until 2019. With respect to government expenditure on education as a percentage of GDP, Kuwait and Oman have reached the target while UAE and Qatar are still below the 4 per cent global target. In the Mashreq, Palestine is the only country that reached the target. In Maghreb countries, Algeria and Tunisia have reached the target.

CHALLENGES MISSING DATA POSE TO EDUCATION SYSTEMS AND TO ACHIEVING SDG 4

The Arab Region has made steady progress to attain SDG 4 especially in some of its subregions like the Mashreq and GCC, yet the region is not on track to achieve the 2030 Agenda. The region suffers from missing quantitative data and insufficient data and analysis for resolving challenges. These have their impact on the education systems of the region and on progress towards achieving the 2030 Agenda. They may have contributed to weak educational systems, low technological productivity and knowledge production, weak science-policy interface, and high youth unemployment despite increased investment in education in some

countries. The proportion of youth not in employment, education, or training (NEET) has increased for both young women and men since at least 2012. The rate in the region was estimated at 34.3 per cent in 2020, compared to a global rate of 22.3 per cent (UNESCO, 2022). The Arab Region is composed of young people (40 per cent) and having these underserved by inadequate education systems will seriously affect their futures. More importantly, the education systems in the Arab Region do not seem to embody the interconnectedness of SDG 4 with other SDGs on human rights, gender equality, sustainability, and ICT – in sum, the foundations of integrated development. This gap is in spite of the fact that the SDGs were designed with their interconnectedness in mind and education can be a central goal in achieving this interconnectivity (UNESCO, 2022). The Arab Region also misses a core tenet of the 2030 Agenda, namely, the concept of "no child left behind" – the principle that sustainable development must include all people regardless of sex, gender, race, migratory status, disability, income, and geographic location – as data on several of these groups was missing.

DATA'S ROLE IN A TRANSFORMATIVE MOVE TO ACCELERATE THE EDUCATION AGENDA 2030

While data is essential for deliberate and transformative change of the education system, it is a tool, not the engine. The Arab Sustainable Development Report 2020 (ESCWA, 2020) highlighted the scale of the challenges facing Arab countries and called for a transformation and dismantling of the structural barriers to achieving the 2030 Agenda. This would include a new vision of education and learning exemplified by curricula aligned with promoting human rights, gender equality, inclusiveness, sustainability, and skills for work. It will require a valued, engaged, and well-trained teaching profession. It will also require information and data that is purposeful for change and adequate for decision-making. I summarize several of the primary data challenges below.

The lack of national standard guiding documents on indicators and data sources poses a challenge to data availability and to identifying missing data gaps. Arab countries have not established common education agendas or monitoring frameworks. Accordingly, it would be useful if countries can develop their national education indicator framework as part of their education sector plan and a guiding document to ensure everyone involved in data collection and analysis will have a single guiding document on concepts and methodologies and avoid misin-

terpretation of information. Only 8 of 22 (36 per cent) Arab countries have updated and approved national strategies for the development of statistics that are disseminated online. Many of the remaining countries still have draft plans under consideration for approval (UNESCO, 2022). UNESCO Beirut facilitated agreement by all the Member States in the region through a series of regional consultations and the regional benchmarks were finalized in July 2021. This was the first step towards the establishment of the national benchmarks by the end of 2021.

Prioritize the production of high-quality, timely, disaggregated data tailored to relevant characteristics in national contexts, especially following the COVID-19 pandemic that has reversed much of the progress achieved to date.

New data sources will have to be identified and the existing ones will have to be assessed, extended, improved, and integrated with other data sources. Filling the gaps requires more effective use of existing data sources, including reliance on household surveys as an important supplement to administrative data collected in schools. Integration with other data sources may warrant better coordination between ministries of education, national statistics offices, and other ministries to increase data flow and avoid duplication of efforts.

Update and modernize statistical legislation and encourage the development of national strategies for the development of statistics (NSDS) as only a third of the region's countries have an updated and disseminated strategy on statistics (ESCWA, 2022).

Provide access to open data, including raw data, to generate knowledge about sustainable development in the region. Build the mechanisms to make use of the data for research and development.

Work to build the necessary statistical capacity and the data analysis capacity to integrate various data sources and the use of data for evidence-informed decision-making.

Attending to the above data challenges will ensure that the underrepresented indicators on education quality, affordability, learning achievement, and missing marginalized groups will be better covered in future education data collection efforts. Without robust data, it is not possible to adequately plan and allocate the resources necessary to ensure that programmes and services achieve their objectives and reach the intended population groups – especially the most underserved. Parallel to that is the need to improve processes and methodologies of policy development and data collection. Only then can the SDG 4 outcomes oriented towards the acquisition of foundational and relevant skills be attained.

REFERENCES

AR-MED III. (2018). Aligning curriculum, teachers and learning assessment to reach SDG 4 targets in Arab countries concept note for third Arab Regional meeting on Education 2030, Dead Sea, Jordan.

ESCWA. (2018). Regional guidebook to improve disability data collection and analysis in the Arab countries: Implementing the Washington Group Questions on functioning. Beirut: United Nations Economic and Social Commission for Western Asia (ESCWA).

ESCWA. (2019). Arab forum for Sustainable Development Report 2019: Empowering people and ensuring inclusiveness and equality in the Arab Region. Beirut: United Nations Economic and Social Commission for Western Asia (ESCWA).

ESCWA. (2020). Arab Sustainable Development Report 2020. Beirut: United Nations Economic and Social Commission for Western Asia (ESCWA).

ESCWA. (2021). Arab forum for Sustainable Development 2021: Accelerating progress on the 2030 Agenda post-COVID. Beirut: United Nations Economic and Social Commission for Western Asia (ESCWA).

ESCWA. (2022). Arab forum for Sustainable Development 2022: Recovery and resilience. Beirut: United Nations Economic and Social Commission for Western Asia (ESCWA).

IDMC. (2022). Global report on internal displacement. Geneva: Norwegian Refugee Council & IDMC.

Luomi, M. (2020). Turning the tide in the Arab Region: How data driven policy can accelerate achieving the SDGs. *Dubai Policy Review*.

Luomi, M., Fuller, G., Dahan, L., Lisboa Båsund, K., de la Mothe Karoubi, E., and Lafortune, G. (2019). Arab Region SDG index and dashboards report 2019. Abu Dhabi and New York: SDG Centre of Excellence for the Arab Region/Emirates Diplomatic Academy and Sustainable Development Solutions Network.

OCHA. (2019). 2019 Humanitarian needs overview: Syrian Arab Republic. New York: OCHA.

UN Economic and Social Council. (2020). Joint report of the Secretary-General, the Washington Group on Disability Statistics and international agencies. New York: Economic and Social Council Statistical Commission.

UNDESA. (2022). World population prospectus 2022. New York: United Nations Department of Economic and Social affairs (UNDESA).

UNESCO. (2022). Progress of SDG 4 in the region: A summary review. Paris: UNESCO.

UIS. (2021). Arab countries Regional report. Bridging SDG 4 and education monitoring in the Arab Region. Montreal: UNESCO Institute for Statistics.

UNHCR. (2022). Population statistics. Geneva: UNHCR.

UNICEF. (2019). *Education*. Amman: Middle East and North Africa Regional Office.

Venkatraman, S. (2021a). SDG 4 progress in Arab Region: Some key highlights. In Fourth Arab Regional meeting on education 2030 (AR-MED IV). Beirut: UNESCO Regional Office.

Venkatraman, S. (September 2021b). Education data quality and collecting data for some challenging SDG 4 targets. In Advanced Preparatory Workshop for National SDG 4 Benchmarking Exercise. Beirut: UNESCO Regional Office.
Wahbe, M. (2021). Can the Arab Region meet the Sustainable Development Goals by 2030? *The Cairo Review of Global Affairs* 40.

6. Priorities for missing data and SDG 4 for countries in Africa

Angela Arnott

INTRODUCTION

This chapter addresses gaps in the education data regime for the Africa region.[1] The SDG monitoring agenda shines a much-needed light on the need for better development data in Africa. However, the global commitment to the SDG 4 targets also creates additional pressure on the already under-resourced capacities of African governments to integrate and monitor their indicators as national priorities, all of which are ideal but some of which are too ambitious.

Despite the reality that over half of the nations in Africa are the poorest in the world, African governments seek to offer quality education services. While some countries have succeeded in achieving universal access, the education challenge facing the continent is that 98 million African children are out of school (UNESCO, 2020a, p. 354), 58 per cent of whom are not in secondary school (UIS, 2019a, 2019b), and 193 million children of primary and lower secondary school age (in 2015) have reached their adolescent years unable to read, write or perform basic numeracy tasks (Watkins, 2013). These education challenges are all the more urgent in the world's youngest continent where almost 60 per cent of sub-Saharan Africa's population is under the age of 25. This scarcity in providing universal access to quality education, given the overwhelming need, must inform any discussion on Africa's data challenges in integrating and monitoring the SDG 4 agenda.

[1] The African Union (AU) countries, the United Nations regions and the Arab states overlap but do so imperfectly. This is addressed in more detail later in this chapter. This chapter specifies when it is referring to AU countries or sub-Saharan Africa.

The targets set by the SDG 4 2030 Agenda and the continent's own African Union (AU) Agenda 2063 (African Union, 2014a) and the Continental Education Strategy for Africa 2016–25 (CESA 16–25) (African Union, 2018) have alerted national and continental policy decision-makers to international best practices in education and training, but the degree to which countries have domesticated these priorities is questionable. Recent evaluation of progress in terms of access to education and achievement of learning outcomes in terms of these targets is recognized as slow in bridging the inequity gaps (African Union, 2022). The recent Transformation of Education Summit (2022) saw African leaders re-commit to transform and revitalize education in Africa.

Reporting on different education agendas was raised as a concern by African Ministries of Education through the 2015 Kigali Statement and later the 2018 Nairobi Declaration (ADEA, 2015; UNESCO, 2018). It is not surprising that data availability has dropped given the unprecedented burdens on national statisticians reporting to both agendas. The CESA 16–25 regional education development agenda has 57 indicators, 17 of which are shared with the 43 SDG 4 indicators. Previously the EFA-MDGs (Education for All-Millennium Development Goals) agenda had only 11 indicators. Recognizing that it is important to simplify, the AU and UNESCO Institute for Statistics (UIS) have agreed to harmonize indicators where they can and focus on piloting a subset of seven benchmark indicators, as a fundamental basis for reporting (AU-IPED, 2022). They will still monitor all the other indicators but use the benchmark indicators as the core of the endline reports on the achievements of targets.

Relative to other regions, African countries have a poor record of reporting on global education indicators. Many African countries provide incomplete data, some more than 2–5 years late and some with methodological divergence from international standard definitions (UNESCO, 2021a). Despite learning being the heart of both the SDG 4 and CESA agendas, fewer than a quarter of countries in Africa have provided the UIS with SDG 4.1 data since 2014 (UIS, 2019b, 2019c).[2]

The reasons for the absence of indicators on the continent are numerous and range from the lack of linkage between data availability and data purpose, insufficient understanding of how such data can inform policy changes, the lack of institutionalization of such processes in national

[2] Author's calculations.

education budgets, the expense and technical complexity requirements outside of available resources, different priorities of Ministries, and lack of data maturity among civil society, amongst others. The real data gap facing Ministries of Education in Africa is operational real-time data on schools, teachers and learners. At a global level there needs to be greater engagement on data needs and purposes as the perspective of African countries has so far been largely absent from global debates on data and urgently requires being heard. There needs to be a greater buy-in globally to the fact that many of the SDG 4 indicators are too numerous, too complex, not financially feasible and not aligned to the priorities of many African Ministries of Education.

Key Sources of Education Data

The UIS, a key custodian of SDG 4 data, has put significant effort into improving their data collection tools as well as pioneering new methodologies for indicator calculations to minimize the phenomenon of missing data and systematic under-reporting of groups and themes. To address national and regional variations, UIS and international organizations apply the ISCED system of education levels classification and use UN Population Development (UNPD) data rather than national population censuses (UISc, 2019, p. 23). At the same time, the building blocks of national statistical systems in Africa remain weak and face challenges in producing the error-adjusted population data needed for the calculation of many SDG indicators. Notwithstanding, many African countries only recognize and use their own national population data as their official statistics, creating potentially conflicting indicators.

National Statistical Offices (NSOs) play a pivotal role in the areas of data collection, coordination, reporting and validation of statistics for the SDGs. Data for several of the SDG 4 indicators rely upon multiple NSO surveys. The challenge with these surveys is that their timing is not synchronized (Demographic and Health Survey (DHS): every five years; Multi Indicator Cluster Survey (MICS): every three years; Labour Force Surveys: annual but erratic) and data definitions are not harmonized, which creates conflicting information. They are no replacement for national information systems collecting comprehensive, annual, large-scale data on schools and learners (UNESCO, 2019) which, as discussed above, have their own limitations for education transformation.

GROUPS AND TYPES OF MISSING DATA

The SDG 4 transformational agenda targets all learners or data groups across a variety of dimensions or data types. Data groups in this chapter refer to subsets of the population or learners who can be constituted as a group based on some common criteria such as age, gender, language, religion and displacement status. Data type refers to a data variable that measures a concept such as data on learning outcomes or child-friendly schools; these largely map to the SDG 4 indicators. Data types intersect with data groups and provide insight into the factors that may lead to groups' increasing stratification by distinguishing their vulnerability further.

In the region, the most significant missing or under-reported data groups are: (a) early childhood learners; (b) youth/adult learners; and (c) disaggregation of learners by different types of vulnerability, for both in- and out-of-school populations. These are critical learner groups to both the SDG 4 and the Africa We Want vision and are close to the fundamental priorities of low-resourced countries seeking to ensure that all learners access quality learning (African Union, 2014b). This chapter will focus on the first two of these three groups due to their clearer alignment with the continent's priorities, their more precise definition, and the overlap of the third group with other chapters in this volume. The importance of data disaggregation and intersectional vulnerabilities are discussed from a global perspective in Chapter 2. The education data gaps and challenges related to the vulnerable groups of refugees and disabled learners are comparable to those discussed in Chapter 3 on Latin America and the Caribbean. The data gaps regarding the important vulnerable group of internally displaced learners in conflict-affected contexts are covered in Chapter 7.

The most significant gaps in terms of data types for the continent include the thematic areas of: (a) learning assessments (SDG 4.1); (b) skills, including numeracy and literacy (SDG 4.6) as well as "relevant skills/ICT skills" (SDG 4.4); and (c) the data gaps on private/non-public/non-formal provision (across multiple SDG indicators). This is not to underrate important discussions related to other SDG targets, subgroups and broader themes, but rather to highlight priority focus areas considered. The third of these will be discussed briefly in the context of the first two missing data groups noted above (i.e., data groups and data types), and many of the challenges regarding data gaps for educational provi-

Table 6.1 Missing Early Childhood Education (ECE) data group

Indicator	Country coverage
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.	26 (48%)
4.2.2 Participation rate in organized learning (one year before the official primary entry age), by sex.	34 (63%)

Source: DHS and UNICEF's Multiple Indicator Cluster Survey (MICS) (UNESCO, 2023). Only 17 African countries use MICS.

sion outside of public, formal schooling in Africa are parallel to those described in Chapter 4 on Asia.

Early Childhood Learners

The first significant under-represented data group is the 0–6-year-old early childhood learners who are monitored by SDG 4.2.1 and 4.2.2. Despite the fact that the (2018) Nairobi Declaration and Call for Action on Education (supported by African ministers) (UNESCO, 2018) made a commitment to "integrated approaches to early childhood development, care and education policies, programming and financing", this data group is underserved by regional and national monitoring frameworks in the region. Mention is made of the importance of early childhood development in CESA, but it lacks a strategic objective and set of monitoring indicators (Table 6.1).

Access to ECE varies widely across African countries and this affects monitoring of early childhood learners, only 28 per cent of whom are registered as receiving ECE support (World Bank, 2021a). In Chad and Mauritania, around one in five children attend organized learning one year before the official primary school age (UNICEF, 2019). Yet several Southern African countries, including Eswatini, South Africa and Zimbabwe, as well as Kenya and Tunisia, are approaching universal pre-primary access. Nevertheless, eight countries spend almost none of their public education budget on pre-primary education (UNESCO, 2021a). African countries are aware that limited access to responsive parenting education, to early stimulation and early learning contribute to poor education outcomes and high dropouts in primary education. However, many countries cannot afford to offer ECE provision. Often, community-run centres fill the gaps. These institutions are frequently

informal and may fall under the mandate of other Ministries. This affects the monitoring of SDG 4.2, particularly as the data regularly excludes private or non-formal, rural or unregistered ECE providers.

Youth and Adult Learners

A second data group systematically under-reported is the education and training of youth and adults. Data is hard to obtain on those in non-formal and/or post-secondary technical and vocational training (TVET) programmes offered outside the public sector. Many of the 58 per cent of youth of secondary school age not enrolled in secondary school are likely to be in some form of employment, apprenticeship or non-formal training (UIS, 2019a). For most African countries, where 85.8 per cent of employment is informal, data on youth and adults participating in non-formal or private sector vocational training programmes is commonly inadequately collected or absent in the region. Another possible source, outside of the SDG 4, is data collected by the the International Labour Organization (ILO) through NSO Labour Force Surveys on SDG 8.6.1 (proportion of youth aged 15–24 years not in education, employment or training (NEET)). These NSO surveys are erratically undertaken by African countries and comparability issues have been raised by the ILO on the harmonization of definitions involving concepts of employment, participation, non-formal and youth. African countries have a poor record of reporting on either set of indicators.

These indicators are particularly difficult to track given the complexity of variables involved in their composition which lend themselves to different definitions and interpretations. Non-formal education (NFE) is defined in ISCED as education that is institutionalized, intentional and planned by an education provider but is considered an addition, alternative and/or a complement to formal education. It may be short in duration and intensity, being typically provided in the form of short courses, workshops or seminars. This type of education largely happens outside of the Ministry of Education and, if offered by governments, are often left out of education data systems[3] because NFE lacks status and is notoriously underfunded.

[3] Sierra Leone has sector-wide Education Management Information Systems (EMIS) but does not include NFE data, despite it being a Ministry directorate.

Table 6.2 Missing data group on youth and adult learners

Indicator	Country coverage	Notes
4.3.1 Participation rate of youth and adults in formal and non-formal education (NFE) and training in the previous 12 months, by sex.	Only 17 (31%) countries provided data between 2010 and 2019.	Coverage is poor as it includes NFE programmes outside of administrative systems. Data is often available for public formal systems. Reliance on DHS restricts data to participating countries every five years.
4.4.1 Proportion of youth and adults with information and communications technology (ICT) skills, by type of skill.	Only 11 (20%) countries, four of which are North African.	
4.6.1 Percentage of population in each age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex.	Data is only available for two countries, South Africa and Ghana reported between 2010 and 2019.	This also relies upon self-reporting or using school grade completion as a proxy for literacy.
8.6.1 proportion of youth aged 15–24 years not in education, employment, or training (NEET).	Data is available for 25 African countries post 2014.	Source is national Labour Force Reports where data is erratically collected.

Source: Author's calculations using public reports from UIS, UNESCO and the AU.

TVET is another subsector which is poorly monitored outside of the public provision in most countries. It is challenging to measure because of the multiplicity of courses offered (largely in the private sector) and the lack of knowledge on the equivalence of programmes in terms of their length, content, accreditation if any, as well as definitional differences. Additionally, on SDG 4.4.1 measuring the level of proficiency in digital literacy, skills critical for the world of work, the ranking is based on individuals' reports that they have carried out selected activities in the previous three months, which may underestimate the percentage of those who possess the skill but have not used it recently. Besides, the official African definition of youth (15–35 years) recognizes that many Africans complete post-secondary education much later than other regions (for which UIS relies on the 15–24 age group). This contextual definition of youth affects the alignment of SDG 4 data with continental policies and priorities more broadly.

Learning Achievement

The global learning crisis has been well documented, but statistics show that it is more widespread in sub-Saharan Africa than in other regions of the world. The region is home to 17 of the 21 countries where more than half of the children lack basic skills, among countries with data (UNESCO, 2020a, 2021a). Despite years of steady growth in enrolment rates, non-proficiency rates in reading and mathematics remain high in the region where 88 per cent of children (202 million) of primary and lower secondary school age were not proficient in reading, and 84 per cent (193 million) were not proficient in mathematics in 2015 (UNSD, 2019).

Most countries in the region struggle to produce the seven SDG 4.1 indicators. The Global Education Monitoring Team estimates that we only know about learning levels at the end of primary education for one-third of the children in the region and only 15 per cent of the school-aged population can be measured on two variables on minimum learning levels (Montjouridès, 2022). They note that the puzzling part, however, is that Africa is perhaps the region where the number of learning assessments has grown the most since the early 2000s. Participating countries in SEACMEQ and PASEC have doubled, and for the 12 countries that are featured in the recent Spotlight report, they found 124 nationally representative assessments at the end of primary carried out between 2000 and 2021. Notably, out of 114 learning assessments in primary carried out between 1995 and 2019 by these 12 countries, only 75 had available reports and only 25 had datasets available. This absence of capacity for data analysis is critically jeopardizing the quality of education monitoring in the region, perhaps even more than the scarcity of data in some instances.

Some 14 African countries have reported data on reading proficiency in the early grades since 2014, and 15 countries have conducted national assessments beyond the 7th year of school. Other non-governmental regional learning assessment tests on literacy and numeracy, such as UWEZO, largely involving East African countries, has grown significantly. It is also notable that in recent years member countries of SEACMEQ have failed to produce national reports which are the basis for the regional comparisons. The challenge is often related to countries' ability to secure funding for regional and international assessments. The impact of COVID-19 added a new element to this situation, forcing

African countries to consider how to monitor learning in the context of remote teaching and learning.

Literacy, Numeracy and ICT Skills

Regarding skills, African countries struggle to provide data on SDG 4.6 on literacy and numeracy proficiency targets. Demographic household surveys collect data on literacy, school attendance and highest levels achieved. They rely upon self-reported data, often using educational attainment data as a proxy for literacy but apply different lengths of school attendance or levels of completion. The contextual utility of this proxy has also been questioned, as a 2015 study on African countries found a weak correlation between educational attainment and literacy (Smith-Greenaway, 2015). Additionally, on SDG 4.4.1 measuring the level of proficiency in digital literacy, skills critical for the world of work, the ranking is based on individuals' reports that they have carried out selected activities in the previous three months, which may underestimate the percentage of those who possess the skill but have not used it recently.

MISSING DATA PURPOSES

In many instances, the data purpose is missing from the SDG 4 indicators for Ministries of Education. They largely fall outside of the normal input indicators Ministries use to inform their budgets and monitor their policy implementation. Despite significant improvements across the continent in new technologies in data collection methodologies and dissemination of data, the outputs of the census survey-based systems are insufficient to address the information needs of decision-makers in the Ministries of Education. Their overwhelming need is for real-time data to inform on operational, managerial and performance issues related to schools, teachers and learners. In terms of a hierarchy of needs, this is the data gap facing Ministries of Education. Their accountability for managing their national systems supersedes the purpose of the SDG 4 targets. The strongest incentive for decision-makers in the Ministries is the supply of data accounting for public expenditure and the delivery of inputs at the national and local levels. Fully funded statistical systems are rare in the region and the added burden to track the SDG indicators is not within the

financial resources of many countries. Indeed, Ministries do not have the financial resources to implement what they know to be best practices.[4]

Missing from the SDG indicator equation is its value to civil society to use the data to hold their government answerable for achieving the SDG targets (Arnott, 2022). Increasingly, a culture of data use is becoming an expected feature of inclusive government sector planning and monitoring, often incentivized by Global Partnership for Education (GPE) funding requirements. In a world of more data, it is essential that citizens can read, work with, analyze and argue with data; but the demand for data by civil society in Africa tends to be low because most lack data literacy, even in the highly IT-enabled societies (such as Ghana, Nigeria, Kenya, Senegal, South Africa) (World Bank, 2021b).

Yet monitoring and reporting on the achievements of a country on the SDG 4 targets also present political risks. Poor national results on comparative regional learning assessments can be viewed as undermining the effectiveness of politicians, particularly during an election year. Data that highlights inequality among groups is not always welcome by national governments. Groups in power may question the data reliability of unpleasant data and worry that drawing attention to disparity gaps in the provision of public services will fuel resentment among the disadvantaged.

For these reasons, the challenges of missing education data, especially the continued invisibility of groups by their absence in monitoring reports, are exacerbated by the gap in understanding the data purpose of the indicators to inform policy and planning (for financial, political and cultural reasons), including data illiteracy by civil society in most African countries.

Relationship of Missing Data to National and Local Systems

Where data groups and education programmes (i.e., data types) are missing from Ministries' priorities, they become invisible. The real power

[4] A low ambition scenario finds that it will cost IDA and IBRD countries US$2.9 billion annually to support the statistical production of measuring the SDGs. A medium ambition scenario finds that it would cost $4.2 billion per year to, in addition, fund capacity-building activities, and in a high ambition scenario, the annual cost is estimated at $5.6 billion to implement all activities under the Cape Town Global Action Plan for Sustainable Development Data (Jolliffe et al., 2021).

of data for development comes from the synergy of sharing data sources, and using it in innovative ways to improve granularity, timeliness and coverage of datasets. Integrating and effectively using data systems involves the engagement of multiple stakeholders – including academic institutions, international organizations and civil society – who are key to enhancing service delivery and policy design. Unfortunately, most African countries are not data mature and lack integrated EMIS database systems. However, this scenario is changing with the recognition by several Ministries of Education of the importance of sector-inclusive database warehouses which allow access to indicator dashboards measuring progress on sector plans at national and local levels (Crouch, 2019; UNESCO, 2021b).

If data is shared in user-friendly formats, education practitioners can learn from past experiences and current challenges to drive future success by allocating financial and human resources more fairly to benefit learners in need (UNICEF, 2022). When communities are provided with integrated, timely, accessible data on the local education situation, then access and learning often improve (UNICEF, 2022). To this end, several Ministries now use school report cards (e.g., Tanzania, The Gambia, Madagascar and Ghana) which are accessible to low-literacy audiences, helping parents, teachers and students to stay informed and hold school managers to account. In a positive example, The Gambia argues that sharing school-level data with communities on their national biannual census-based learning assessments has resulted in significant regional improvements.[5]

Information feedback to the schools is critical for learning adjustment. Although research on the usability of learning assessment data as a diagnostic tool to improve classroom teaching and learning has yet to be confirmed (UNESCO, 2021b), by combining the results with other sector databases and geographic locations, decision-makers can identify challenges for redress. A recent survey of 42 African countries found that the use of learning assessment results falls into three main categories: curriculum/policy reform (26 countries); teacher professional development (29 countries); and targeting a specific area of learning (25 countries) (UNESCO, 2020). However, the lack of a culture of data use was demonstrated by the fact that a third of countries did not answer this

[5] KIX Education Symposium (October 2022) The Gambia presentation, Data informs Learning Achievement. Addis Ababa.

question. A recent analysis on the South African Learning Assessment data indicates that the irrelevance of data derived from these tests discourages teachers from trusting the data. Teachers often need additional support to promote a culture of evidence-informed decision-making in their classrooms (Kanjee and Moloi, 2014).

Data Integration Nationally, Regionally and Globally

Despite the hopes that SDG 4 data will inform the transformation of education for the public good, data integration nationally, regionally and globally is often misaligned for various reasons.

Firstly, the national and the global education data on the same indicator and year can look different and are not recognizable as having a common source. Although all education data is extracted from national official sources, in order to allow for cross-country comparability, the UIS is obliged to use ISCED classifications of education levels and use UN population data instead of the national official population census data. The UN population data provides a different denominator which affects the calculation of ratios (e.g., Gross Enrolment Ratios (GER)). This can cause conflict with national authorities where development partners prefer to base their financial support on the UIS data.[6]

Secondly, even where national data exists, it is not available in UIS databases. Several countries have a poor record in releasing their national data to the UIS, and some with delays of between two to five years. For various reasons – a lack of data consistency with previous years, or inaccuracy or incompleteness, or definitional differences or the country is not submitting any data to UIS – the data is not published. The enrolment data of four of the most populated six countries in Africa are not comprehensively covered by UIS databases, namely, South Africa (2018), Nigeria (2018), Kenya (2016) and Ethiopia (2015), but more recent data is available in these countries' national EMIS (UIS, 2021c).

Thirdly, civil society and the media seem unaware of their role in monitoring the SDG 4 indicators. Except for comparisons of national learning achievement against other African countries for media interest, national decision-makers seldom refer to SDG 4 indicators as performance benchmarks. Even country comparisons reported by African regional economic

[6] Author's experience of working in the Ministry of Education, Ghana in 2007 where GER rates were different.

communities and the AU remain largely ignored by national education stakeholders and the media.

Africa also faces the challenge that, at the global level, data on the continent has been separated into two regions (48 sub-Saharan African countries and six North African countries in a cluster with Western Asia), instead of a continent of 54 African countries.[7] This adds to challenges of the AU and Regional Economic Communities in reconciling regional estimates of achievement and comprehensive accounting of continental progress on the SDG 4 objectives. On the face of it, this segregation of international data may support the international development community but lacks direct resonance at continental and regional levels.

CONCLUSION

Missing data groups and types affect the development agenda of the Africa region. It is difficult to make a comprehensive assessment on regional achievement on the SDG 4 targets given the paucity of data on key data groups (especially early childhood learners and youth) and on critical data types (particularly learning achievements and skills). This is compounded by the mismatch between the SDG 4 data regime and countries' need for data to inform policy formulation, rather than primarily to monitor implementation. There are a number of obstacles and disincentives for African countries to invest scarce resources in monitoring many of the SDG 4 and CESA targets. Coming from a low base of data monitoring and use, many countries lack the skills and capacity to address the monitoring requirements; lack the systems for effective data use; have little social pressure from their populations to measure the progress on achieving these targets; and have more pressing national priorities to address, which are fundamental to social services serving the needs of the majority. This does not diminish the ambitious ideals of the SDG 4 and CESA targets, but it needs to be recognized that monitoring many of these is beyond the scope of many countries in this region, and thus there is a need to reduce and simplify these indicators.

[7] Although the AU recognizes 55 countries, including Sahrawi Arab Democratic Republic as a country, this report is limited to 54 countries as it uses UN data for its analysis.

BIBLIOGRAPHY

ACPF, & AMC. (2019). *In the firing line: The war on Africa's Children*. African Child Policy Forum (ACPF) & The AfricaWide Movement for Children (AMC). https://app.box.com/s/1llqndggbgpvjesbllirnc38xx9no7co Accessed 7 November 2021.

ADEA. (2015). Sub-Saharan Africa Regional Ministerial Conference on Education Post-2015 (Kigali Statement). Abidjan: Association for the Development of Education in Africa (ADEA).

African Union. (2014a). *Key documents of Agenda 2063*. Addis Ababa: African Union Commission. https://au.int/en/documents/20141012/key-documents-agenda2063. Accessed 7 November 2021.

African Union. (2014b). *The Africa we want*. Addis Ababa: African Union Commission. https://au.int/en/agenda2063/overview. Accessed 7 November 2021.

African Union. (2018). *Continental Education Strategy for Africa (CESA 16–25) Indicators Manual*. Addis Ababa: African Union Commission.

African Union. (2022). *Draft concept note on the AU theme of 2022: As the year of nutrition*. Addis Ababa: African Union Commission. https://au.int/sites/default/files/documents/41426-doc-EX_CL_1288_XXXIX_E.pdf. Accessed 1 February 2023.

Anderson, K., & Muskin, J. (2018). *Learning champions: How 15 countries, cities, and provinces came together to rethink learning assessment*. Washington, DC: The Center for Universal Education (CUE), Brookings.

Arnott, A. (2022, 14 April). Using data for accountability. *NORRAG Missing Data Blog Series*. https://www.norrag.org/using-data-for-accountability/. Accessed 1 February 2023.

AU-IPED. (2022). *The AU IPED Strategy 2023–2025*. Tunis: African Union Pan-African Institute for Education for Development (AU-IPED).

Care, E. (2020). *Optimizing assessment for all: Assessment as a stimulus for scaling 21st century skills in education systems*. Washington, DC: The Center for Universal Education (CUE), Brookings.

Centre for Global Development, & The African Population and Health Research Centre. (2014). *Delivering on a data revolution in Sub-Saharan Africa*. Washington, DC: Center for Global Development. https://www.cgdev.org/sites/default/files/CGD14-01%20complete%20for%20web%200710.pdf. Accessed 7 November 2021.

Choi, J., Mark, D., & Zainab, U. (2020). *The future of work in Africa: Harnessing the potential of digital technologies for all.* Washington, DC: World Bank.

CREATE. (2010). *Seeds of their struggle: The features of under and overage enrolment among Grade 4 learners in South Africa*. Brighton: CREATE, Centre for International Education.

Crouch, L. (2019). *Meeting the data challenge in education*. Washington, DC: Global Partnership for Education. 2019-07-15-kix-data-final-english.pdf (globalpartnership.org).

ED, & NCES. (1991). *Standards for education data collection and reporting (SEDCAR)*. Washington, DC: US Department of Education (ED) & National Centre for Education Statistics (NCES).

European Union, & United Nations. (2018). *Technical report on statistics of internally displaced persons current practice and recommendations for improvement*. Luxembourg: Publications Office of the European Union. https://ec.europa.eu/eurostat/documents/3859598/9316015/KS-GQ-18-003-EN-N.pdf.

Figlio, D., & Özek, U. (2019). Unwelcome guests? The effects of refugees on the educational outcomes of incumbent students. *Journal of Labor Economics*, *37*(4), 1061–96. https://doi.org/10.1086/703116

Glassman, A., & Sandefur, J. (2014). *Why African stats are often wrong*. Washington: DC and London: Center for Global Development. https://www.cgdev.org/blog/why-african-stats-are-often-wrong.

Global Disability Rights Now. (2021). *Disability in Kenya*. Eugene, OR: Mobility International. https://miusa.globaldisabilityrightsnow.org/infographic/disability-kenya/

IGAD. (2017). *Djibouti plan of action on refugee education in the IGAD member states*. Djibouti: Intergovernmental Authority on Development.

ILO. (2018). *Women and men in the informal economy: A statistical picture (third edition)*. Geneva: International Labour Organization.

Johnston, D. (2014). *Benefits and costs of the data for development targets for the post-2015 development agenda*. Tewksbury, MA: Copenhagen Consensus Center. http://www.post2015consensus.com. Accessed 7 November 2021.

Jolliffe, D., Kilic, T., Mahler, D., & Randolph, P. (2021). *How much should governments spend on data?* Washington, DC: World Bank. https://blogs.worldbank.org/opendata/how-much-should-governments-spend-data. Accessed 7 November 2021.

Kanjee, A., & Moloi, M. (2014). South African teachers' use of national assessment data. *South African Journal of Childhood Education*, *4*(2), 90–113. https://doi.org/10.4102/sajce.v4i2.206.

Lewin, K.M. (2020). Beyond business as usual: Aid and financing education in Sub Saharan Africa. *International Journal of Educational Development*, *78*, 102247. https://doi.org/10.1016/j.ijedudev.2020.102247.

Mbuen, E. (2011). *Implementing the UN Convention on the Rights of Persons with Disabilities in Africa: Opportunities and challenges*. Budapest and London: Mental Disability Advocacy Centre. https://www.mdac.org/en/implementing-un-convention-rights-persons-disabili. Accessed 7 November 2021.

Montjouridès. (2022). *GEM Spotlight Panel. AU IPED Strategy*. Tunis: AU-IPED.

Orton, M., & Rowlingson, K. (2007). *Public attitudes to economic inequality*. York: Joseph Rowntree Foundation.

Oxford Policy Management. (2021). *Recommendations to strengthen Uganda's Education Response Plan for Refugees and Host Communities (ERP)*. Oxford: Oxford Policy Management. https://www.opml.co.uk/files/Publications/A2241-maintains/recommendations-to-strengthen-uganda-s-education-response-plan-for-refugees-and-host-communities-finance-edited.pdf?noredirect=1. Accessed 7 November 2021.

Over, M. (2014, March). *A "data revolution" of the people, by the people, and for the people – not just for advocates*. Washington, DC and London: Center for Global Development.

RTI. (2015). *Status of early grade reading in Sub-Saharan Africa*. Washington, DC: US Agency for International Development (USAID).

Sandefur, J., & Glassman, A. (2014). *The political economy of bad data: Evidence from African survey & administrative statistics*. Washington, DC and London: Center for Global Development.

Smith-Greenaway, E. (2015). Educational attainment and adult literacy: A descriptive account of 31 Sub-Saharan Africa countries. *Demographic Research*, *33*(35), 1015–34.

Statistics Sierra Leone. (2018). *Sierra Leone Integrated Household Survey*. Freetown: Statistics Sierra Leone. https://www.statistics.sl/images/StatisticsSL/Documents/SLIHS2018/SLIHS_2018_New/sierra_leone_integrated_household_survey2018_report.pdf. Accessed 7 November 2021.

UIS. (2018). *Quick guide to education indicators for SDG 4*. Montreal: UNESCO Institute for Statistics.

UIS. (2019a). *Fact Sheet no. 56*. Montreal: UNESCO Institute for Statistics. https://uis.unesco.org/sites/default/files/documents/new-methodology-shows-258-million-children-adolescents-and-youth-are-out-school.pdf. Accessed 7 November 2021.

UIS. (2019b). *UIS education data release: September 2019*. Montreal: UNESCO Institute for Statistics. https://uis.unesco.org/sites/default/files/documents/ip59-uis-education-data-release-september-2019.pdf. Accessed 7 November 2021.

UIS. (2019c). *SDG 4 Data Digest: How to produce and use the global and thematic education indicators*. Montreal: UNESCO Institute for Statistics. https://uis.unesco.org/sites/default/files/documents/sdg4-data-digest-2019-en_0.pdf. Accessed 7 November 2021.

UIS. (2021a). *Guidelines for data collection to measure SDG 4.7.4 and 4.7.5*. Montreal: UNESCO Institute for Statistics. https://uis.unesco.org/sites/default/files/documents/guidelines_sdg_474_475_final_v3.pdf. Accessed 1 February 2023.

UIS. (2021b). *Summary report of Africa regional technical consultations on regional benchmarks for SDG 4 and CESA 16-2*. Montreal: UNESCO Institute for Statistics. https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2021/08/Final-Outcomes-Report_-Regional-benchmarks_Africa_20210729.pdf. Accessed 7 November 2021.

UIS. (2021c). *UIS statistics*. Montreal: UNESCO Institute for Statistics.

UNDESA. (2020). *World social report 2020: Inequality in a rapidly changing world*. United Nations Department of Social and Economic Affairs. New York: UNDESA.

UNESCO. (2018). *Nairobi Declaration and Call for Action on Education. Bridging continental and global education frameworks for the Africa We Want*. Paris: UNESCO. https://en.unesco.org/sites/default/files/2018-04-23_en_pace_2018_working_documents_2304_rev.pdf. Accessed 7 November 2021.

UNESCO. (2019). *Global Education Monitoring Report 2019: Migration, displacement and education: Building bridges, not walls*. Paris: UNESCO.
UNESCO. (2020a). *Global Education Monitoring Report 2020: Inclusion and education: All means all.* Paris: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000373718.
UNESCO. (2020b). *Learning assessments in Sub-Saharan Africa. Teaching and learning educators network for transformation*. Paris: UNESCO.
UNESCO. (2021a). *Progress Report on SDG4 Targets by SADC member states*. Harare: UNESCO.
UNESCO. (2021b). *KIX Africa 19 Data Challenge Seminar Report Three. Feedback mechanisms for schools*. Harare: UNESCO.
UNESCO. (2023). *World Inequality Database on Education*. https://www.education-inequalities.org/. Accessed 6 September 2023.
UNICEF. (2019). *The State of the World's Children 2019. Children, food and nutrition: Growing well in a changing world*. New York: UNICEF. https://www.unicef.org/media/106506/file/The%20State%20of%20the%20World%E2%80%99s%20Children%202019.pdf. Accessed 7 November 2021.
UNICEF. (2021). *MICS6 TOOLS*. New York: UNICEF. https://mics.unicef.org/tools. Accessed 7 November 2021.
UNICEF. (2022). *Data must speak: Country initiatives: Explore how this UNICEF education initiative works around the world.* New York: UNICEF. https://www.unicef.org/education/data-must-speak-country-initiatives
United Nations. (2017). *Cape Town Global Action Plan for Sustainable Development Data*. New York: United Nations. https://unstats.un.org/sdgs/hlg/Cape_Town_Global_Action_Plan_for_Sustainable_Development_Data.pdf"https://unstats.un.org/sdgs/hlg/Cape_Town_Global_Action_Plan_for_Sustainable_Development_Data.pdf. Accessed 7 November 2021.
United Nations. (2019). *Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*. New York: United Nations.
UNSD. (2019). SDG progressing: SDG4 quality education. New York: United Nations Statistics Division (UNSD). https://unstats.un.org/sdgs/report/2019/goal-04/. Accessed 7 November 2021.
UNSD. (2022). *SDG indicator metadata (Harmonized metadata template – Format version 1.1)*. New York: United Nations Statistics Division (UNSD). https://unstats.un.org/sdgs/metadata/files/Metadata-10-06-01.pdf. Accessed 1 February 2023.
van der Berg, S. (2008). How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation*, *34*(3), 145–54. https://doi.org/10.1016/j.stueduc.2008.07.005
Watkins, K. (2013). *Too little access, not enough learning: Africa's twin deficit in education*. Washington, DC: The Center for Universal Education (CUE), Brookings. https://www.brookings.edu/opinions/too-little-access-not-enough-learning-africas-twin-deficit-in-education/. Accessed 7 November 2021.
WHO, & World Bank. (2011). *World report on disability 2011*. Geneva: World Health Organization (WHO). https://apps.who.int/iris/handle/10665/44575. Accessed 7 November 2021.

World Bank. (2020). *Global Estimate of Children in Monetary Poverty: An Update, Poverty & Equity*. Washington, DC: World Bank.
World Bank. (2021a). Data. Washington, DC: World Bank. https://datatopics.worldbank.org/education/. Accessed 7 November 2021.
World Bank. (2021b). *The World Bank Development Report 2021: Data for better lives*. The World Bank. https://www.worldbank.org/en/publication/wdr2021. Accessed 21 February 2022
World Bank. (2022). *School enrollment, preprimary (% gross) – Sub-Saharan Africa*. Washington, DC: World Bank. https://data.worldbank.org/indicator/SE.PRE.ENRR?locations=ZG. Accessed 1 February 2023.
World Population Review. (2022). *Poorest countries in Africa 2022*. Walnut, CA: World Population Review. https://worldpopulationreview.com/country-rankings/poorest-countries-in-africa. Accessed 21 February 2022.

7. Missing education data on internally displaced people (IDPs)

Chiara Valenti and Louisa Yasukawa

INTRODUCTION

There were 71.1 million people living in internal displacement worldwide at the end of 2022 due to conflict, violence, and disasters, the highest figure on record (IDMC, 2023a). The exact number of children among them is unknown, but there are estimated to be about 11.6 million internally displaced people (IDPs) aged between 5 and 11, and 9 million between 12 and 17 (IDMC, 2023b).

Like all children, those internally displaced have a right to education at all stages of their journey (OCHA, 1998). Target 4.5 of Sustainable Development Goal 4 (SDG 4) specifies the need to ensure access to education for children living in vulnerable situations, which is widely recognised to include IDPs (UNHCR, 2020). In times of crisis, however, children displaced within borders are among those most neglected. They often experience significant disruptions to learning and lose access to formal schools entirely (IDMC, 2019). Financial barriers, loss of civil documentation, distance, language barriers, and safety concerns are just some of the obstacles IDPs encounter in accessing quality education (IDMC, 2022b). Displaced girls, children from Indigenous groups, and those living with a disability tend to face heightened education challenges (Gakunga & Gathinye, 2020). The lack of comprehensive, reliable, and comparable data on IDP's education amplifies their invisibility and reinforces educational barriers.

National governments are primarily responsible for the provision of education to IDPs and related data collection. There is also growing recognition that disaggregating SDG 4 indicators by displacement status is essential to fulfil the SDGs' commitment to "leave no one behind" (IISD, 2018). There are, however, few examples of systematic national monitor-

ing of the issue. Although humanitarian and development organisations also collect data on IDP's education, the definitions and methodologies they use vary depending on their purpose, which hampers the compilation of national-level figures.

In addition, lack of awareness of existing data sources and inadequate coordination between data collectors means that the data collected on IDP's education is often not utilised (Shephard et al., 2021). Yet, such data is crucial for understanding the scale of the issue, planning and costing effective responses, measuring progress, and evaluating the impact of education interventions.

This chapter explores the data landscape on IDP's education. It begins by providing an overview of some of the main data gaps, sources, and challenges. It then discusses how conceptual challenges, political factors, capacity and resource constraints, and a lack of coordination between data collectors contribute to these challenges. It concludes by outlining ways forward to safely address existing data gaps and highlights promising tools and initiatives to strengthen data quality, sharing, and interoperability.

As conflict, violence, and disasters continue to uproot people from their homes, including IDPs in education data is essential to designing better programmes and addressing the consequences of their plight on their current and future lives.

DATA GAPS

Data on Internally Displaced Children in General

Internally displaced children are invisible in data for two reasons. Firstly, compared with asylum seekers, refugees, and migrants, IDPs of all ages are largely unaccounted for. National governments are primarily responsible for the data collection on internal displacement, but many lack the capacity or will to do so systematically. Governments may be reluctant to even recognise IDPs, particularly if they have contributed to displacement.

As IDPs are not captured in data collected at border crossings, their presence is generally only recorded if they receive assistance or are registered by relevant authorities in displacement camps, evacuation centres, or collective shelters. Those that find refuge in host communities or informal settlements remain largely untraced.

Monitoring displacement is easier in countries that have national registries of IDPs, such as Colombia and Georgia. In most countries affected by internal displacement, however, such national monitoring of the phenomenon does not exist. The Internal Displacement Monitoring Centre (IDMC) is the official global source of data on people internally displaced due to conflict, violence, and disasters. It verifies, curates, and triangulates data reported by a range of partners and sources, including the UN and other international organisations, the media, civil society organisations, and research institutions (IDMC, n.d.-a). On this basis, IDMC works to publish data that is as geographically comprehensive and comparable as possible, but availability varies considerably within and between countries.

Secondly, internally displaced children are largely missing in national and international data because the limited data that is available on IDPs is rarely disaggregated by age. Only about 5 per cent of the displacement records IDMC collected in 2021 included some form of age disaggregation (IDMC, 2022a). Time and resource constraints posed during displacement crises mean that data is often collected at the household level. In addition, humanitarian organisations regularly gather information on "crisis-affected children" or "children in need", without a further breakdown of their displacement status. The extent to which collecting data on displaced children is possible or advisable is also limited, rightfully, by ethical considerations.

IDMC estimates the number of school-aged IDPs at the national and international level based on national age distribution data from the UN World Population Prospects, but their true number remains unknown. For example, IDMC estimated that there were 2.2 million IDPs aged between 5 and 17 in Afghanistan at the end of 2022 (IDMC, 2023b).

It is difficult to compile accurate figures on the number of school-aged IDPs, their location and needs at a given point in time because of the dynamic nature of IDP's movements, conceptual challenges, and variations in the way data collectors define IDPs and the age groups they use.

Data on IDP's Education

While timely and reliable information on the number of school-aged IDPs is limited, comprehensive data on their education is even harder to come by. This is even though disaggregating data by displacement status is relevant to several indicator frameworks on education, such as the SDG 4 indicator framework, among others.

The Inter-agency Network for Education in Emergencies' (INEE) minimum standards are intended to guide authorities, humanitarians, and other stakeholders in the delivery of quality education in emergency settings, including displacement crises (INEE, 2021). The related indicator framework notes that information on attendance rates, learning outcomes, and children's safety and others should be disaggregated by displacement status when appropriate.

Despite these calls, information on IDP's enrolment, attendance, and completion rates in formal education are not consistently collected at the national level, and out-of-school rates for IDPs are unavailable. The number of IDPs participating in vocational training or informal schooling is also unknown.

Comprehensive information on the quality of education that IDPs receive and their literacy, numeracy, and other learning outcomes is even more scarce. While useful tools to gather data on academic learning, socio-emotional learning, and the psychosocial well-being of students have been developed, they are rarely applied to contexts of displacement. The lack of disaggregated data on these issues also limits our understanding of how the education impacts of displacement vary depending on a child's sex, age, ethnicity, disability status, and other characteristics.

There is a significant gap when it comes to data on teachers in displacement contexts and they tend to be overlooked in global education indicators relating to crisis settings (Mendenhall et al., 2019). Information on the availability of teachers, their daily attendance and their level of training is highly fragmented, but is crucial to informing resource allocation, identifying gaps, and assessing the quality of education internally displaced children receive. Further information on teachers' profiles, background, psychosocial well-being and professional development needs can assist in tailoring training programmes and informing interventions aimed at increasing the supply of teachers in underserviced areas (Mendenhall et al., 2018). At the same time, better information on the number of IDPs of a school-age would assist in mobilising sufficient funding for teacher recruitment and development.

Costing data on actual and projected education expenditure is limited in general, and even more so for IDPs. Several methodologies have been developed to estimate the cost of providing children with an education based on publicly available information, including refugee children and those affected by crises (ODI, 2016; UNESCO, 2015; World Bank & UNHCR, 2021). Few costing exercises have focused specifically and exclusively on internally displaced children, however (IDMC, 2022c).

Data on the funds that governments, humanitarian, and development actors dedicate to IDP's education can inform strategic planning and resource mobilisation.

The significant gaps in data at the national level mean internationally comparable figures on IDP's education are unavailable. UNESCO's Institute for Statistics (UIS) is the official source of internationally comparable education data more generally and is used to measure progress towards SDG 4 (UIS, n.d.). Although it contains data for more than 200 countries from a range of sources, none of its indicators are currently disaggregated by displacement status. UNESCO's Section of Education for Migration, Displacement and Emergencies intends to standardise data and information on a collectively agreed set of priority indicators, including for IDPs. Although for now, internationally comparable figures on their school attendance and completion, out-of-school rates, and learning outcomes are lacking.

DATA SOURCES

This section examines the potential government and humanitarian sources of data on IDP's education before discussing the factors that constrain the effective collection and use of such information.

National Education Management Information Systems

National education management information systems (EMIS) have the potential to be an important source of data on IDP's education. EMIS collect, integrate, process, and disseminate data and information to support decision-making, planning, monitoring, and management at all levels of an education system (van Wyk & Crouch, 2022). A comprehensive EMIS gathers information on school enrolment, attendance, completion rates, learning assessments, student health, finance, teacher characteristics, and administrative statistics from a range of data sources, such as annual school censuses.

Several countries have adapted their systems to identify migrants and refugees among their student populations and capture better data on their education status. Education stakeholders in Syria have added an indicator on IDPs to annual school censuses as part of a broader transformation of the country's EMIS (UNESCO, 2021). In certain areas of the country, an online integrated management information system for schools has been piloted to collect more comprehensive and up-to-date data on

learner access, attendance rate, and performance. As this system also collects data on students' movements, this can be used as a proxy for displacement status. While such an example is promising, few countries have the financial and technical resources to develop more dynamic and operational EMIS. As a result, there is generally no distinction between IDPs and other students in administrative data.

IDP Registries

National registries of IDPs can be used to gather information on IDP's age and education status. They can also be connected to education ministry data systems or statistical offices to track IDP's access to education. Yet, successful examples of this in practice are rare. One exception is Colombia's victims' registry, which has been integrated with information from the education and other ministries using individuals' national ID (Shephard et al., 2021).

Censuses and Surveys

Another potential source of data on IDP's education are censuses and other national population-based surveys, such as demographic health surveys (DHSs) and multi indicator cluster surveys (MICSs). Iraq's 2018 MICS, for example, disaggregated households by displacement status and contained questions on school attendance rates and out-of-school children (Central Statistical Organization et al., 2019). Another example is Colombia's 2015 DHS, which looked at school attendance rates, repetition and drop-outs, and disaggregated households by those who had moved internally as a result of violence, which was used as a proxy for displacement status (Ministerio de Salud de Colombia, 2015). Such assessments are costly and resource-intensive, however, and political sensitivities and methodological challenges can limit the type of data they can gather.

Data Collected by Humanitarian Organisations

Considering the gaps in national monitoring of IDP education, international organisations, such as the UN Office for the Coordination of Humanitarian Affairs (OCHA), the International Organization for Migration (IOM), IMPACT Initiatives/REACH, the Joint IDP Profiling Service (JIPS), and IDMC collect data on IDP's education. They employ

a range of methods including key informant interviews, focus group discussions, and household surveys, each with its own strengths and weaknesses.

IOM's multi-sectoral location assessments (MSLAs), for example, provide annual high-level snapshots on education access on a relatively wide scale for different countries. The 2021 MSLA for Mozambique, for example, assessed 69 sites across three different provinces hosting 201,689 IDPs (IOM, 2022). It included indicators on the number of displaced children with access to education facilities, the percentage of functioning schools in the sites, main barriers to education, and distance to the nearest education facility.

REACH's multi-sector/multi-cluster needs assessments (MSNAs/MCNAs) are requested annually by OCHA to inform humanitarian needs overviews (HNOs) and response plans. They generally aim for nationwide coverage and can be used to gather information on the status of education facilities and household education needs. More than 12,500 IDPs, returnees, and non-displaced households were surveyed for the 2021 MCNA for Iraq, which included indicators on the percentage of households with a child not attending school regularly, disaggregated by displacement status and location site (REACH, 2022).

The joint education needs assessments (JENAs) coordinated by education clusters and the profiling exercises conducted by JIPS are other important sources of data but are conducted less regularly and in fewer countries than the assessments conducted by IOM and REACH. The 2019 JENA for out-of-school children in north-west Syria provided information on the number of out-of-school IDPs, their school attendance prior to displacement, reasons for dropping out, and main barriers to education (Education Cluster, 2020). Similar information was gathered through the profiling exercise conducted by JIPs of 4,662 displaced and non-displaced households in Sittwe township, Rakhine state in Myanmar in 2016 and 2017 (JIPS, 2017).

OCHA's HNOs are another key source of information on the proportion of IDPs in need of education support. They use the joint intersectoral analysis framework and bring together information from secondary data analysis and assessments conducted by REACH, IOM, and education clusters.

IDMC also collects primary data on IDP's education using an original household tool and key informant interviews. This approach has been used to gather information on the number of school-aged children in education, main barriers, satisfaction levels, and cost in specific locations

as part of assessments on the socioeconomic impacts of displacement in Colombia, Ethiopia, Indonesia Kenya, Nepal, Nigeria, Papua New Guinea, Somalia and Vanuatu (IDMC, n.d.-b).

FACTORS INFLUENCING THE AVAILABILITY AND USE OF DATA

The data collected by governments and humanitarian organisations has the potential to provide valuable insights on IDP's education. There are several reasons why the availability and use of quality, current, and comparable data remains insufficient, however.

Conceptual and Methodological Challenges

Although the UN Guiding Principles on Internal Displacement provide a definition of an IDP, data collectors vary in their approaches to defining an IDP, the causes of displacement, and the criteria they use to determine the end of displacement (Shephard et al., 2021). As a result, there is no standardised approach when it comes to deciding who would be classified as an IDP in education data.

Children born into displacement pose additional conceptual questions. A strict interpretation of the Guiding Principles on Internal Displacement suggests that children born after at least one of their parents has been displaced are not by definition IDPs, given that they themselves were not forced to leave their homes. The Expert Group on Refugee and IDP Statistics (EGRIS) also advises against including them in the definition because it would increase the overall count of IDPs even if no new displacements had taken place (EGRIS, 2020).

It recommends instead that children born into displacement be counted as a separate "IDP-related" population group. In practice, most states either do not count them as IDPs or do not specify, though there are some exceptions. In Georgia, for instance, an underage person is entitled to an IDP status if one or both parents have and/or had IDP status (Government of Georgia, 2014).

The International Recommendations on Internally Displaced Persons Statistics (IRIS) published by EGRIS provide authoritative guidance on how to define and safely identify IDPs in censuses, surveys, and administrative data and registries (EGRIS, 2020). Complementing EGRIS's work, JIPS has published recommendations intended to better identify IDPs when conducting surveys (JIPS, 2021). Despite these important

resources, a standardised and consistent approach to identifying IDPs in data is still lacking in practice.

Political Factors

Political factors can hinder the effective collection and use of data on IDP's education (Shephard et al., 2021). Governments may not be willing to recognise IDPs and identify them in their data systems, particularly if they are a perpetrator of violence. In such contexts, it may be more appropriate for humanitarian actors to lead data collection on the issue. Equally, some IDPs may choose not to be identified in education data if doing so increases their risk of physical violence, exclusion, discrimination, or stigmatisation.

As such, while disaggregating data by displacement status would help bridge crucial data gaps, it must be carefully weighed against any risk of harm. Where it is not safe or politically sensitive to ask about displacement status directly, information on students' movements or locations can serve as useful proxies. Unique school IDs can also be used to ensure personally identifiable information is not shared between agencies.

Capacity and Resource Constraints

Despite the importance of evidence-based programming, government and humanitarian actors responding to complex crises often have limited resources and funding to collect education data in general, let alone on IDPs (Montjouridès, 2013; Shephard et al., 2021). EMIS may not be fully functioning in crisis-affected countries. EMIS that rely on data from annual school censuses are often insufficient in highly fluid displacement crises, but few countries have the financial and technical resources to develop more dynamic and operational data systems.

National household surveys are an effective way of gathering representative information on IDP's education, but they are also the most costly and resource-intensive method. Governments and humanitarian organisations rarely have the resources to conduct them at regular intervals.

While international donors dedicate funds to supporting the education of IDPs, they tend to fund programmes that run in parallel to national education systems. Investments in data collection activities are often directed towards monitoring and evaluating the programmes they fund, rather than national education systems.

Lack of Coordination and Data Sharing

Lack of coordination and data sharing between different data collectors can hamper efforts to monitor IDP's education and means the data that is collected is not effectively utilised. For example, although Ukraine has an IDP registry, IDPs are not obliged to provide an IDP certificate when their child enters school (Shevtsova & Fitisova, 2022). Instead, they only have to show it if they receive some benefit or support for their education. As a result, IDP's school enrolment is not monitored at the local level, which means it cannot be generalised at the national level. If the data systems of the registry and the education ministry were harmonised, authorities would be able to identify IDPs among school students more easily and track their attendance.

Local and international organisations, academics, and others who collect data on IDP's education tend to do so using different methodologies, indicators, and school-age groups. This reduces the interoperability of the data, which makes it difficult to standardise and compare figures. Even when national surveys are carried out and are disaggregated by displacement status, IDPs living in inaccessible areas or those who have lost documentation may not be included, leading to uneven and incomplete data.

WAYS FORWARD TO ADDRESS DATA CHALLENGES

Having identified the existing data sources and challenges when it comes to IDP's education, this following section outlines why it is important to collect such data. It then identifies ways in which these gaps may be addressed to strengthen evidence-based policymaking.

The Importance of Identifying IDPs in Education Data

In calling for improvements in the availability and quality of data on IDP's education, it is important to consider why such data is necessary and what role it will serve. This raises the question as to why it is useful to identify IDPs in education data, rather than simply improving education data systems overall. Although many children in crisis-affected communities require education support, evidence suggests that internally displaced children are often disproportionately affected in crisis settings and are at greater risk of exclusion (Save the Children, 2020; UNICEF,

2020). Lower enrolment rates for displaced children than among their non-displaced counterparts have been documented in assessments in sub-Saharan Africa, Latin America, and the Asia Pacific region (IDMC, 2021, 2022a; World Bank Group, 2019).

Despite sharing many of the same vulnerabilities as children in host communities and other crisis-affected groups, internally displaced children tend to face distinct challenges relating to their displacement that can give rise to specific education needs. Their separation from livelihoods, assets and networks, loss of civil documentation, and geographical and social marginalisation are all common impacts of displacement, which can affect IDP's education in unique ways. Understanding these differentiated impacts and education disparities between crisis-affected groups can assist government, humanitarian and development agencies to design more inclusive policies and ensure the children most at risk of missing out are not left behind.

At the same time, unlike refugees, who are afforded special protection under the 1951 Refugee Convention, IDPs remain at the mercy of their domestic jurisdictions, which compounds their invisibility (UNHCR, 1951). Improving data and evidence on the impacts of displacement on education can raise awareness of IDP's plight and highlight the cost of inaction. Strengthening data collection efforts can therefore encourage nationally owned action on internal displacement and enable governments to track progress more effectively (UN High-Level Panel on Internal Displacement, 2021).

In addition to tracking progress on the SDGs and other frameworks, data on IDP's education is key for monitoring IDP's progress towards achieving durable solutions to displacement. The Inter-Agency Framework on Durable Solutions for Internally Displaced Persons notes that restoring children's right to education is a prerequisite for their achievement of durable solutions (IASC, 2010). Indicators of progress towards achieving a durable solution include the percentage of IDPs with access to adequate and quality primary education at the minimum and the absence of legal or administrative obstacles preventing IDPs from going to school. The lack of data on these indicators impedes accurate monitoring.

Although it is important to improve education data systems in crisis settings in general, such factors highlight the value of identifying IDPs in education data, where it is possible to do so safely and in accordance with their rights.

Improving Data on Internally Displaced Children in General

The availability and quality of data on IDPs overall must be improved if IDP's education is to be monitored systematically. There is a general recognition across the international community that it is the role of national governments to generate and use reliable internal displacement data.

Disaggregating data on IDPs by age is crucial to identifying internally displaced children. Further disaggregation by sex, ethnicity, disability status, and other characteristics is also necessary to identify their intersecting vulnerabilities and tailor programmes more effectively. This need for more systematic data disaggregation is recognised in the SDG target 17.18. EGRIS also makes recommendations for internal-displacement data to be disaggregated by age, sex and date and place of birth, as well as the number of times they have been displaced, the date of their first and most recent displacement, the main reason for their movements, their place of habitual and current residence, whether their parents have also been displaced, and type of habitation (EGRIS, 2020).

Existing tools and guidance could facilitate these efforts. For example, UNICEF and the Washington Group on Disability Statistics have developed a module for inclusion in surveys to identify children with disabilities and disaggregate data accordingly (UNICEF & Washington Group, 2017).

Addressing Gaps in Education Data on IDPs

Finally, comprehensive, comparable, and disaggregated data on displaced children's education is indispensable to improve planning and overcome barriers to learning. In addition to qualitative and quantitative data on IDP's access to school, information on the quality of education they receive, their out-of-school and completion rates are also needed.. Information on teacher availability, as well as their training and professional development needs is essential to assessing the quality of education delivered and improving teacher supply and retention. Data on education expenditure on IDPs and the financial cost of providing them with quality education would assist in informing budgeting and resource mobilisation. Further research is needed on the longer-term consequences of missed education on IDP's development and future opportunities, and the impacts of displacement on the education of non-displaced children.

There are promising efforts to address these gaps. South Sudan and Syria show how the development of a more operational and dynamic

EMIS can be used to gather data on IDP's education. The Children on the Move project of IOM's displacement tracking matrix (DTM) has produced guidelines and tools to facilitate collaboration between the DTM and education clusters in gathering and using data that partners need.

Save the Children has developed a toolkit to help measure gaps in children's achievement of durable solutions, which includes indicators related to their education (Save the Children, 2019). Cambridge Education provides technical advice and support for governments and schools in many countries affected by internal displacement to measure learning outcomes and improve the quality of education services and programmes (Cambridge Education, n.d.).

Strengthening Data Quality, Sharing, and Interoperability

For this data to effectively address these gaps it must also be standardised, safely and ethically collected, safely disseminated, and regularly updated. The International Data Alliance for Children on the Move (IDAC), which is supported by UNICEF, IOM, UNHCR, the Organisation for Economic Co-operation and Development, and Eurostat, is a cross-sectoral global coalition comprised of governments, non-governmental organisations, think tanks, academics, and civil society. The initiative is intended to foster streamlined approaches to improving statistics and data on displaced children and has established a dedicated working group on IDPs (UNICEF, 2022). Similarly, the Responsible Data for Children (RD4C) initiative is a joint endeavour between UNICEF and The GovLab at New York University that aims to produce guidance, tools, and leadership to support governments and humanitarian organisations to carry out responsible and ethical data management (RD4C, n.d.).

Several other examples demonstrate a growing awareness of the need to strengthen data quality, sharing, and interoperability in this field. INEE has established a Data Reference Group on Education in Emergencies, which brings together dozens of organisations that work on related topics to share their data, methodologies, approaches, and experiences (INEE, 2020).

UNESCO's Section of Education for Migration, Displacement and Emergencies is developing a global data portal that will include information on IDPs. It is also working to strengthen institutional information systems for data-driven education in emergencies and crises and implementing country-specific interventions (UNESCO, 2021, 2022). UNICEF is drawing on IDMC's displacement datasets and risk model-

ling and other existing information to refine data on displacement associated with climate change by analysing the number of children displaced, their needs, and how to address them (UNICEF & IDMC, 2022).

While such examples are promising, some experts have noted that the proliferation of separate initiatives led by different organisations is splintering efforts to address data gaps on IDP's education, leading to disparate ways of working and duplication (Shephard et al., 2021). This has prompted calls for a single multi-lateral agency to lead efforts in coordinating and harmonising data collection efforts.

CONCLUSION

The report of the UN Secretary General's High-Level Panel on Internal Displacement (2021) and the subsequent action agenda highlights the key role that access to quality education plays in achieving durable solutions. The UN Guiding Principles on Internal Displacement and the regional frameworks and national policies that incorporate them also recognise access to education and learning as a basic right of displaced children (OCHA, 1998).

Education helps to foster IDP's integration and strengthen social cohesion. When designed and delivered effectively, it constitutes a powerful tool for reducing conflict and fragility, and with it the risk of displacement. To take effective action to unlock these benefits and advance solutions to internal displacement, it is essential to improve the availability of timely, reliable, and comparable data on IDP's education.

To address the lack of qualitative and quantitative information that continues to hamper the design and implementation of tailored policies and programmes, more investment in local data collection, and in national and regional monitoring systems is needed. Information should be collected and used in ways that prioritise children's safety and agency, and consider the diversity of their experiences.

REFERENCES

Cambridge Education. (n.d.). *Providing Advice and Support*. Retrieved 5 December 2022, from https://www.camb-ed.com/intdev/advice

Central Statistical Organization et al. (2019). *Iraq: Multiple Indicator Cluster Survey*. https://mics-surveys-prod.s3.amazonaws.com/MICS6/Middle%20East%20and%20North%20Africa/Iraq/2018/Survey%20findings/Iraq%202018%20MICS%20Survey%20Findings%20Report%20%5B20211213%5D_English.pdf

Education Cluster. (2020). *Joint Education Needs Assessment for Out of School Children (JENA) – Syrian Arab Republic*. https://reliefweb.int/report/syrian-arab-republic/joint-education-needs-assessment-out-school-children-jena-december-2019

EGRIS. (2020). *International Recommendations on Internally Displaced Persons Statistics (IRIS)*. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-005

Gakunga, D.K., & Gathinye, J.M. (2020). Educational Challenges That Girls of School Going Age Face in Their Quest for Education in IDP Camps: The Case of Maai-Mahiu IDP Camp, Nakuru County, Kenya (pp. 225–38). Brill. https://doi"https://doi.org/10.1163/9789004417014_013

Government of Georgia. (2014). Law of 2014 on Internally Displaced Persons – Persecuted from the Occupied Territories of Georgia. Refworld. https://www.refworld.org/docid/44ab85324.html

IASC. (2010). *IASC Framework on Durable Solutions for Internally Displaced Persons*. https://interagencystandingcommittee.org/other/iasc-framework-durable-solutions-internally-displaced-persons

IDMC. (n.d.-a). *Global Internal Displacement Database*. IDMC. Retrieved 5 December 2022, from https://www.internal-displacement.org/database

IDMC. (n.d.-b). *Socioeconomic Impacts of Internal Displacement*. IDMC. Retrieved 27 November 2022, from https://www.internal-displacement.org/research-areas/socioeconomic-impacts-of-internal-displacement

IDMC. (2019). *Twice Invisible: Accounting for Internally Displaced Children*. https://www.internal-displacement.org/sites/default/files/publications/documents/201911-twice-invisible-internally-displaced-children.pdf

IDMC. (2021). *Impacts of Displacement: Conflict and Violence in Quibdo and Caucasia, Colombia.* https://www.internal-displacement.org/sites/default/files/publications/documents/220302_IDMC_SocioeconomicImpacts_Colombia.pdf

IDMC. (2022a). *Disaster Displacement: Vanuatu Country Briefing*. https://www.internal-displacement.org/publications/disaster-displacement-vanuatu-country-briefing

IDMC. (2022b). Informing Better Access to Education for IDPs. https://www.internal-displacement.org/sites/default/files/publications/documents/221114_IDMC_Informing-Better-Access-to-education-for-IDPs.pdf

IDMC. (2023a). *2023 Global Report on Internal Displacement*. https://www.internal-displacement.org/sites/default/files/publications/documents/IDMC_GRID_2023_Global_Report_on_Internal_Displacement_LR.pdf

IDMC. (2023b). Global Internal Displacement Database. www.internal-displacement.org/database/displacement-data

IISD. (2018, 11 December). *To Leave No One Behind, Brief Calls for Considering IDPs in SDG Implementation*. https://sdg.iisd.org:443/news/to-leave-no-one-behind-brief-calls-for-considering-idps-in-sdg-implementation/

INEE. (2020). New INEE Reference Group to Drive Reforms and Set Global Standards for EiE Data. https://inee.org/blog/new-inee-reference-group-drive-reforms-and-set-global-standards-eie-data

INEE. (2021). *INEE Minimum Standards Indicator Framework*. https://inee.org/resources/inee-ms-indicator-framework

IOM. (2022). *IOM Mozambique: Multi-Sectoral Location Assessment Round 7, Mozambique*. https://reliefweb.int/report/mozambique/iom-mozambique-multi-sectoral-location-assessment-round-7-mozambique-cabo-delgado

JIPS. (2017). *Sittwe Camp Profiling Report*. https://www.jips.org/uploads/2018/11/Sittwe_Camp_Profiling_Report_LQ.pdf

JIPS. (2021). *Identifying Internally Displaced Persons in Surveys: Approaches and Recommended Questionnaire Modules*. https://www.jips.org/jips-publication/identifying-internally-displaced-persons-in-surveys-approaches-and-recommended-questionnaire-modules-jips-2021/

Mendenhall, M., Gomez, S., & Varni, E. (2018). Teaching amidst Conflict and Displacement: Persistent Challenges and Promising Practices for Refugee, Internally Displaced and National Teachers. *Paper commissioned for the 2019 Global Education Monitoring Report, Migration, displacement and education: Building bridges,not walls*. https://unesdoc.unesco.org/ark:/48223/pf0000266060

Mendenhall, M., Gomez, S., Varni, E., & Guven, O. (2019). Teachers in Forced Displacement Contexts: Persistent Challenges and Promising Practices in Teacher Supply, Quality, and Well-being. *Journal on Education in Emergencies*, 5(1), 123–32. https://doi.org/10.33682/98cy-cb43

Ministerio de Salud de Colombia. (2015). *Encuesta Nacional de Demografia y Salud Tomo 1*. https://dhsprogram.com/pubs/pdf/FR334/FR334.pdf

Montjouridès, P. (2013). Education Data in Conflict-Affected Countries: The Fifth Failure? *Prospects*, 43(1), 85–105.

OCHA. (1998). Guiding Principles on Internal Displacement.

ODI. (2016). *A Common Platform for Education in Emergencies and Protracted Crises*. https://cdn.odi.org/media/documents/A_common_platform_for_education_in_emergencies_and_protracted_crises_evidence_paper.pdf

RD4C. (n.d.). *Responsible Data for Children*. Retrieved 5 December 2022, from https://rd4c.org/

REACH. (2022). *IRAQ: MCNA IX Brief Multi-Cluster Needs Assessment*. https://reliefweb.int/report/iraq/iraq-mcna-ix-brief-multi-cluster-needs-assessment-january-2022

Save the Children. (2019). *Durable Solutions for Children Toolkit*. https://resourcecentre.savethechildren.net/document/durable-solutions-children-toolkit/

Save the Children. (2020). *Submission to the High-Level Panel on Internal Displacement by Save the Children*. https://www.un.org/internal-displacement-panel/sites/www.un.org.internal-displacement-panel/files/published_save_the_children_submission.pdf

Shephard, D., Buckner, E., Smiley, A., & Hatch, R. (2021). Mapping the Ecosystem of Education Data for Internally Displaced Persons in the Middle East and Beyond: Issues, Challenges, and Recommendations. https://pdf.usaid.gov/pdf_docs/PA00ZH7D.pdf

Shevtsova, O., & Fitisova, A. (2022, 8 April). Education Data in Displacement Context: What Ukraine Does and Has to Do to Collect Data on IDPs? *NORRAG*.

https://www.norrag.org/education-data-in-displacement-context-what-ukraine-does-and-has-to-do-to-collect-data-on-idps/

UIS. (n.d.). *Data for the Sustainable Development Goals*. Retrieved 5 December 2022, from https://uis.unesco.org/

UN High-Level Panel on Internal Displacement. (2021, September). *Shining a Light on Internal Displacement: A Vision for the Future*. https://internaldisplacement-panel.org/

UNESCO. (2015). Pricing the Right to Education: The Cost of Reaching New Targets by 2030. https://unesdoc.unesco.org/ark:/48223/pf0000232197

UNESCO. (2021). Strengthening Education Management Information Systems (EMIS) and Data for Increased Resilience to Crisis: Country Case Study: Syria. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376064

UNESCO. (2022). *UNESCO Supports Ethiopia in Enhancing the Production, Harmonization and Use of Education in Emergency (EiE) Data*. UNESCO. https://en.unesco.org/news/unesco-supports-ethiopia-enhancing-production-harmonization-and-use-education-emergency-eie

UNHCR. (1951). *Convention and Protocol Relating to the Status of Refugees*. https://www.unhcr.org/protection/basic/3b66c2aa10/convention-protocol-relating-status-refugees.html

UNHCR. (2020). Data disaggregation of SDG Indicators by Forced Displacement. https://www.unhcr.org/5fd237b84.pdf

UNICEF. (2020). *UNICEF's Submission to the High-Level Panel on Internal Displacement*. https://www.un.org/internal-displacement-panel/sites/www.un.org.internal-displacement-panel/files/published_unicef_submission.pdf

UNICEF. (2022). *International Data Alliance for Children on the Move (IDAC)*. https://data.unicef.org/resources/international-data-alliance-for-children-on-the-move/

UNICEF, & IDMC. (2022, 22 August). Leveraging Data to Unlock New Solutions for Children Uprooted by Climate Change. Patrick J. McGovern Foundation. https://medium.com/patrick-j-mcgovern-foundation/leveraging-data-to-unlock-new-solutions-for-children-uprooted-by-climate-change-44abc9cf59ba

UNICEF, & Washington Group. (2017). *Disability Statistics Module on Child Functioning*. https://data.unicef.org/resources/module-child-functioning-concept-note/

van Wyk, C., & Crouch, L. (2022). *Efficiency and Effectiveness in Choosing and Using an EMIS*. UNESCO Institute for Statistics.

World Bank Group. (2019). *Informing Durable Solutions for Internal Displacement in Nigeria, Somalia, South Sudan, and Sudan*. https://documents1.worldbank.org/curated/en/761091557465113541/pdf/Volume-A-Executive-Summary.pdf

World Bank, & UNHCR. (2021). *The Global Cost of Inclusive Refugee Education*. World Bank. https://doi.org/10.1596/35238

8. Indigenous data sovereignty and missing education data

Jacob Prehn, Karen Martin and Gawaian Bodkin-Andrews

INTRODUCTION AND BACKGROUND

Globally, nation states' education systems generally prioritise the dominant ethnic groups' worldviews within education and its components: curriculum, measures of student success, and data (Anderson et al., 2016; Lopez, 2021; Suina & Chosa, 2021). This dominant ethnic emphasis (e.g., whiteness) on educational data is particularly problematic for both Indigenous education (the education of Aboriginal and Torres Strait Islander peoples and international First Nations peoples) and Indigenous Studies (e.g., the teaching of Aboriginal and Torres Strait Islander and international First Nations cultures, communities, histories, sciences, lived experiences). That is, representative and meaningful measures of practices and outcomes that are meant to signify 'quality' Indigenous education and Indigenous Studies, namely, good Indigenous data, is largely overlooked by education systems and their institutions, researchers, and policymakers (Lopez, 2021; Suina & Chosa, 2021; Walter, Kukutai et al., 2021).

We argue that in Australia (and all CANZUS countries: Canada, Australia, Aotearoa, New Zealand, and the United States), good Indigenous educational data is largely missing because Indigenous peoples are absent, or minimally represented within the data itself. This prioritisation of the dominant ethnic group within educational data is ultimately to the detriment of minority ethnic groups, such as Aboriginal and

Torres Strait Islander and international First Nations Peoples[1] because their worldviews become marginalised, are not seen as valuable, and are erased (Anderson et al., 2016; Behrendt et al., 2012; Walter, 2016; Walter, Carroll et al., 2021). Further, this marginalisation of Indigenous worldviews also works against majority ethnic group student development because they receive a restricted and narrower educational experience that does not expose them to multiple worldviews (Hart, 2010; Kaomea, 2009).

Within the Anglo-colonised CANZUS countries, the marginalisation of Indigenous worldviews by the state and non-state educational systems is an ongoing issue (Kukutai & Cormack, 2021; Lopez, 2021; Lovett, Jones, & Maher, 2021; Suina & Chosa, 2021). For example, in Australia, scholars have explored how Indigenous worldviews (inclusive of epistemology, ontology, and axiology) can be included within curricula and measures of educational success in primary (Bishop, Vass, & Thompson, 2021; Martin, 2017; Williamson-Kefu, 2022), secondary (Donovan, 2015; Ober et al., 2022), and tertiary education systems (Behrendt et al., 2012; Martin, 2008; Nakata, 2010; Page, Trudgett, & Bodkin-Andrews, 2019; Prehn et al., 2020).

In Australia, over the last two decades, there has been an increasing effort to rectify the shortcomings of the primary, secondary, and tertiary education systems by decolonising and Indigenising curricula and education structures, and the training and employment of Indigenous professional and academic staff (Price, 2012; Trudgett, Page, & Coates, 2022; Universities Australia, 2017). To achieve this, measures have included greater recruitment and retention of Indigenous teachers (Andersen, O'Dowd, & Gower, 2015; Universities Australia, 2017), the employment of Aboriginal Education Workers (AEWs) (Peacock & Prehn, 2019; Price et al., 2017), decolonising and Indigenising the curriculum (Bodkin-Andrews et al., 2021; Hart et al., 2012; Nakata, 2010; Page et al., 2019; Prehn et al., 2020), and Indigenous cultural activities and pro-

[1] From this point on, Aboriginal and Torres Strait Islander and Indigenous international First Nations Peoples will be referred to as Indigenous peoples. With the Australian Anglo-colonised context Aboriginal is an aggregated descriptor for many unique Indigenous peoples with their own distinct identity, cultural practices, customs, lore, and histories (Dudgeon et al., 2014). Similarly, the Torres Strait Islander peoples consist of five distinct peoples whose lands range from the top of the mainland Australian content, stretching almost to Papua New Guinea (Dudgeon et al., 2014).

grammes (Bodkin-Andrews et al., 2013; Harwood et al., 2015; Martin, 2017).

As already suggested though, despite efforts to decolonise all levels of education within Australia, representative Indigenous education and Indigenous studies data is still largely absent. The result of this ignorance of Indigenous peoples and our worldviews results in what Palawa sociologist, Distinguished Professor Maggie Walter (2018) describes as the 5Ds of Indigenous data: disparity, deprivation, disadvantage, dysfunction, and difference. For example, within this 5D data narrative, Indigenous peoples are measured against the European majority population, and these simplistic binary measurements generally position Indigenous peoples as being in need of non-Indigenous salvation (Walter & Carroll, 2021).

The aim for Aboriginal and Torres Strait Islander populations to reach the same outcomes as the non-Indigenous population continues to ignore their cultural differences in aspirations and life values, and results in data that are focused on difference, disparity, disadvantage, dysfunction, and deprivation (Lovett, Jones, & Maher, 2021, p. 44).

At present, Australian educational data is largely missing Indigenous worldviews and Indigenous input at each stage of the data lifecycle, instead Anglo-Australians and their cultural norms mostly control the entirety of Indigenous education data systems, including analysis, dissemination, and subsequent policy interpretations thereof (Bodkin-Andrews et al., 2017; Walter & Carroll, 2021). The educational data lifecycle begins at the conceptualisation phase, and includes the 'who, what, when, why, and how' of data. It is at this stage, where Indigenous input is often overlooked, or at best Indigenous advice is saught, but rarely acted upon. Further, the other stages of the data lifecycle such as creation and collection, through to analysis, writing up, dissemination, and ongoing project sustainability are also often missing Indigenous peoples' involvement (Maiam nayri Wingara, 2018; Walter, 2016).

To rectify the issue of missing Indigenous involvement in educational data, we argue that the concepts of Indigenous Data Sovereignty (IDSov) and Indigenous Data Governance (IDGov) are key mechanisms for Indigenous people to attain good education data inclusive of their worldview (Kukutai & Taylor, 2016; Lopez, 2021; Walter, Kukutai et al., 2021). IDSov and IDGov present an opportunity for Indigenous worldviews and educational priorities to be appropriately woven into the composition of educational data, and for Indigenous educational pri-

orities to be recognised alongside the non-Indigenous focus, rather than Indigenous worldviews being assimilated (Kukutai & Cormack, 2021).

ABORIGINAL PEOPLES AND TORRES STRAIT ISLANDER PEOPLES: CONTEXTS

Aboriginal and Torres Strait Islander peoples have lived on what is now known as the Australian continent for time immemorial (e.g., we refuse to submit to ever-changing scientific measurements based on non-Indigenous measures of 'time'), and at the point of British Invasion in 1788, it was estimated that there were over 1,000,000 peoples with up to 250 different language groups and over 800 dialects (AIATSIS, 2022; Dudgeon et al., 2014; Ryan, 2012). With the current Indigenous population estimated to be 984,000 peoples (ABS, 2022), it is important to note these numbers are still not equivalent to those prior to British Invasion. In addition, despite ongoing efforts to revive Indigenous languages, it is estimated that only 120 of these are still spoken today, with 90 per cent being judged as endangered (AIATSIS, 2022). From this, it must be understood that British Invasion and subsequent colonisation, with frontier wars, massacres, disease, destruction, and theft of Countries, enforced poverty, and the Stolen Generations has led many seminal Indigenous scholars to argue that British 'settlement' was, and still is, a blatant act of epistemic, cultural, and physical genocide (Behrendt, 2001; Rigney, 1999). Many more scholars have noted that the forces of colonisation (and even genocide) are still prevalent today in the ongoing marginalisation and oppression of Indigenous peoples through government policies and practices across the likes of health and wellbeing, law, and education (Dudgeon et al., 2014; Martin, 2008; Nakata, 2010; Paradies, 2016; Watson, 2009).

PRIMARY AND SECONDARY EDUCATION IN AUSTRALIA

In Australia, the provision of primary and secondary education (and its funding) is predominantly the constitutional responsibility of nine State and Territory Governments (Australian Government Productivity Commission, 2022). Ideally, the policy orientated decision making of these governments should be evidenced based, but this chapter argues that such 'evidence' is not representative of Indigenous peoples. Data on compulsory schooling is collected though an annual National Schools

Statistics Collection managed by the nine Australian state and territory education departments (Australian Curriculum, Assessment and Reporting Authority, 2020). The data pertain to both government schools (government, education departments) and non-government, and independent/private schooling providers. Schooling census data is held by each of the nine jurisdictions (disaggregated to region and school levels) and submitted to the Australian Curriculum, Assessment and Reporting Authority (ACARA), an independent statutory authority. This is the main architecture for national schooling data and measures the goals and objectives of the Australian Education Ministers' Council (Australian Curriculum, Assessment and Reporting Authority, 2020).

While this data now exists, it was not until the late 1900s that targeted data recognition of Indigenous students began to be collected (e.g., National Aboriginal Education Policy, 1989, 1995). Today, although the more recent Aboriginal and Torres Strait Islander Education Strategy 2015 remains current, it has somewhat been usurped by the education target of the Council of Australian Governments' (COAG) 'Closing the Gap on Indigenous Disadvantage' strategy (2008).

In 2022, the now refreshed and renamed 'Closing the Gap Strategy' still only has two education targets specifically relating to primary and secondary school (National Indigenous Australians Agency, n.d.), and only one involves 'supporting indicators' for school attendance and retention rates, literacy and numeracy results, and PISA (Programme for International Student Assessment) test for 15-year-old students (Australian Government Productivity Commission, 2022). Further, within PISA, the only way to identify Indigenous peoples is through language, however, as detailed above in the Australian context, Indigenous languages are 90 per cent endangered and often not readily spoken, resulting in Indigenous people not being accurately represented within the PISA data.

The implications are that the responsibility for setting and measuring the goals, indicators, and outcomes for the education of Aboriginal and Torres Strait Islander students has the following strong limitations:

- Remains entrusted to jurisdictions (states and territories);
- Remains underpinned by an ideology where there have been no national education policy changes since 2015;
- Does not receive the same attention as other socio-economic indicators in this Closing the Gap Strategy; and

- Is exacerbated by missing educational indicators and so, missing educational data.

Further within the Closing the Gap Strategy, there are four Priority Reforms, of which Priority Reform number four is 'Shared Access to Data and Information at a Regional Level'. The aim of this Priority Reform is that:

> Aboriginal and Torres Strait Islander people have access to, and the capability to use, locally-relevant data and information to set and monitor the implementation of efforts to close the gap, their priorities, and drive their own development. (Closing the Gap, 2022, p. 1)

At present, work is being undertaken by Australian governments and Aboriginal and Torres Strait Islander peoples to achieve this priority reform, so its achievement remains to be seen. Unfortunately, it may be argued that broader government approaches to Indigenous education (e.g., Closing the Gap) have largely failed to acknowledge the potential for 'data' and educational programmes that is not only created from Indigenous epistemic foundations, but also highly valued (and used) by Indigenous students, families, and communities (Martin, 2017). Instead, it has been repeatedly argued that successive governments have committed to a form of ideological settler violence where education has been the tool of Indigenous student, family, and community erasure. Where non-Indigenous and Western educational norms and measures are the dominant, and too often only, visible outcome in the non-Indigenous 'Indigenous' education data. As powerfully argued by Gumbaynggirr scholar Lilly Brown (2019, p. 66), the likes of Closing the Gap narratives are currently little more than 'research and policy premised on the a priori assumption that the problem of Indigenous people is first and foremost disadvantage and deficiency'. That is such conclusions of ongoing 'disadvantage and deficiency' are more reflective of the individual, systemic, and epistemic racisms embedded within government analyses (and subsequent policies) of Australia's education systems than the capabilities and potential of Indigenous students themselves (Bodkin-Andrews et al., 2021; Moodie, Maxwell, & Ruldolf, 2019). Measuring Indigenous children with data that is not necessarily reflective of their Indigenous worldviews is problematic and harmful. Data issues such as missing Indigenous worldviews and priorities within state educational data used in the Closing the Gap Strategy reiterates the importance of IDSov and IDGov to be operationalised (Lovett, Jones, & Maher, 2021).

SUSTAINABILITY DEVELOPMENT GOAL (SDG) 4: QUALITY EDUCATION

At a global level, the sovereignty and rights of Indigenous peoples far exceed the United Nations Sustainability Development Goals (SDGs) such as SDG 4: Quality Education. For example, as stated in the United Nations Declaration on the Rights of Indigenous Peoples (United Nations, 2008), Indigenous peoples ought to:

- Have the right to establish and control their education systems, attain state education without discrimination, and get an education in their own culture and language (Article 14);
- Have the right to dignity and diversity of their cultures, traditions, histories, and aspirations which shall be appropriately reflected in education and public information (Article 15); and
- Have the right to the improvement of their socio-economic conditions inclusive of education (Article 21).

Although SDG target 4.5 aims to 'Eliminate all discrimination in education' by 2030, and identifies Indigenous peoples as a specific marginalised group, several of the current data practices used by nation states to report on the progression of SDG 4 assimilate Indigenous educational outcomes into aggregated nation state data (SDG Tracker, 2022). We argue that IDSov and IDGov are a meaningful and respectful method of appropriately attaining data to measure how Indigenous peoples globally are progressing to achieve SDG 4, a quality education that aligns with their rights as outlined by the UNRDIP (2008).

INDIGENOUS DATA SOVEREIGNTY (IDSOV)

IDSov is a global movement regarding the rights of Indigenous peoples to have ownership, control, access, and possession of data relating to their lives (First Nations Information Governance Centre, 2014). The IDSov movement started in the 1990s with work by the Canadian Steering Committee of the First Nations Regional Longitudinal Health Survey (Schnarch, 2004). Their push to have sovereignty over their data was a 'political response to colonialism and the role of knowledge production in reproducing colonial relations' (Espey, 2002, p. 1). Then, in 1998, the First Nations Information Governance Centre (2014, p. 1) established the

OCAP® Principles which stand for: Ownership, Control, Access, and Possession.

While the OCAP® Principles were conceptualised by Canadian First Nations people, the historical and ongoing experiences of colonisation are similar for other Indigenous peoples across the globe (Anderson et al., 2016). As a result, Australia, Aotearoa (New Zealand), the United States and other Indigenous peoples globally have defined their own IDSov principles and protocols and have progressed work on their operationalisation. In these countries, it is important for non-Indigenous research organisations, researchers, policymakers, and governments to understand how Indigenous peoples are progressing their IDSov movements, and this includes Indigenist and Indigenous education data.

In 2018, the Australian IDSov collective Maiam nayri Wingara (MnW) held their inaugural Indigenous Data summit (Maiam nayri Wingara, 2021). There participants defined Indigenous Data, Indigenous Data Sovereignty, and Indigenous Data Governance in an Australian context (Maiam nayri Wingara, 2018). Additionally, five Indigenous Data Governance principles were developed, to exert the rights of Aboriginal and Torres Strait Islander peoples in relation to their data. The principles are as follows:

- Exercise control of the data ecosystem including creation, development, stewardship, analysis, dissemination and infrastructure;
- Data that is contextual and disaggregated (available and accessible at individual, community and First Nations levels);
- Data that is relevant and empowers sustainable self-determination and effective self-governance;
- Data structures that are accountable to Indigenous peoples and First Nations; and
- Data that is protective and respects our individual and collective interests.

This initiated the Australian IDSov movement, and the process has begun to operationalise IDSov in various contexts (e.g., community and government, non-governmental organisations).

In 2019, collectively Indigenous peoples from the CANZUS countries and other locations around the globe, including Europe, Africa, and Latin America, developed the CARE (Collective Benefit, Authority to Control, Responsibility, and Ethics) Principles for good Indigenous Data Governance (RDA IIDSIG, 2022). The CARE Principles are an

Indigenous global response to the movement towards open data and open sciences. They have been developed to sit alongside the non-Indigenous FAIR (Findable, Accessible, Interoperable, and Reusable) Principles which, by themselves, can contribute to the ongoing marginalisation of Indigenous peoples and their data (Carroll et al., 2020). Together, the Indigenous CARE Principles along with the non-Indigenous FAIR Principles promote more equitable participation in the processes of data governance and reuse for Indigenous peoples.

INDIGENOUS DATA GOVERNANCE (IDGOV)

The concept of IDSov is enacted through IDGov (Lovett et al., 2019; Rainie et al., 2017; Smith, 2016; Walter & Carroll, 2021). The notion of IDGov is Indigenous peoples having power and authority over the ownership, control, access, and possession of their data (i.e., the OCAP® Principles) (First Nations Information Governance Centre, 2014). The concept of IDGov has two key aspects (Carroll, Rodriguez-Lonebear, & Martinez, 2019; Hudson et al., 2017; Walter & Carroll, 2021):

- The governance of data (controlling access and the use of Indigenous data); and
- Data for governance (to achieve Indigenous community aspirations).

To achieve IDGov, strong Indigenous leadership is key because Indigenous led and controlled decision making ensures that Indigenous worldviews (inclusive of epistemology, ontology, and axiology), along with priorities, values, and cultures are embedded within the data (Smith, 2016; Walter & Carroll, 2021). Strong Indigenous leadership is needed throughout the data lifecycle, and the missing data project amongst other projects, entities, and policymakers have a role to play in growing, maintaining, and enhancing the potential for Indigenous leadership to occur.

IDGov needs to occur across the whole Indigenous data lifecycle. This means commencing at the conceptualisation phase and continuing through to development and data collection, to analysis and dissemination (Rainie et al., 2017). Therefore, governance over Indigenous data is not just about stewardship, but collecting data which is relevant and needed by Indigenous peoples to achieve their needs and aspiration. Largely, while Indigenous communities/nations rely on external data collected by government institutions, large philanthropic bodies, and a wide diversity of consultative bodies, these data often fail to reflect community

needs, priorities, and aspirations. This imbalance risks commitment to self-determination, limits informed policymaking decisions, and restricts Indigenous progress. As noted by Smith (2016, p. 130):

> Strong governance creates checks and balances to ensure that data collection supports the priorities of a group or organisation, implements agreed standards for data quality control and works to ensure data are available in a timely way. Ineffective governance of data can lead to uninformed decision-making, low participation by membership, project failures, loss of reputation and credibility, and missed development opportunities.

The second aspect of IDGov is Indigenous peoples having the data they need for self-governance. Data for governance recognises Indigenous community aspirations to aid in nation (re)building. When Indigenous people are the decision makers, Indigenous Nations and community representatives can harness capacity and implement strategic decisions about their own affairs, and make a comprehensive effort to (re)build and enhance their governance structures (Hudson et al., 2017; Smith, 2016). Secondly, by implementing Indigenous data for governance, it empowers the community to support its members' development and aspirations (Hudson et al., 2017; Lovett et al., 2019).

For Indigenous peoples to successfully (re)achieve self-determination and autonomy, having data to support successful governance is crucial. The process of strengthening and rebuilding data for governance is a challenging journey for many Indigenous peoples. However, strong Indigenous leadership and IDSov (Walter & Carroll, 2021) are key components to ensure any priorities and aspirations of the Indigenous peoples are being met and incorporated into the data unlike the current 'missing data' phenomena.

THE UNITED NATIONS DECLARATION ON THE RIGHTS OF INDIGENOUS PEOPLES (UNDRIP)

The IDSov movement uses the UNDRIP as a mechanism for Indigenous peoples to assert their rights to their data and their right to education (United Nations, 2008). The UNDRIP resolution was passed in 2007 after a vast majority of the 159 countries, in total 144 countries, voted in favour, 11 abstained, and 4 voted against. The four states (Canada, Australia, Aotearoa/New Zealand, and the United States) that voted against the resolution were the CANZUS countries, highlighting the ongoing tensions that exist in these Anglo-colonised nations. Over the

following years, due to social pressures, the four CANZUS countries eventually reversed their position on the UNDRIP and now support it.

The UNDRIP contains 46 Articles which articulate individual and collective minimum standards of Indigenous rights, including education, cultural expression, identity, language, employment, health, and other areas (United Nations, 2008). Many of the 46 Articles have some component that is interconnected with education. Articles 18–23: are rights that enable improvement of Indigenous socio-economic conditions in areas such education and training, employment, housing, sanitation, health, and social security (Davis, 2016). For example, Article 21.1 states (United Nations, 2008, p. 17):

> Indigenous peoples have the right, without discrimination, to the improvement of their economic and social conditions, including, inter alia, in the areas of education, employment, vocational training and retraining, housing, sanitation, health and social security.

This Article demonstrates how Indigenous peoples have inherent rights to assert Indigenous interests in relation to Indigenous data and governance of that data, to improve their education and socio-economic positioning. Moreover Article 18 states that (United Nations, 2008, pp. 15–16):

> Indigenous people have the right to participate in decision-making in matters which would affect their rights, through representatives chosen by themselves in accordance with their own procedures, as well as to maintain and develop their own Indigenous decision-making institutions.

At the centre of IDSov is Indigenous decision making, therefore this Article illustrates that Indigenous peoples have a right to participate in *all* matters pertaining to their data at an individual or collective level (Carroll et al., 2020; Lovett et al., 2019). This should also include Australian governments' meaningful commit to educational standards also stipulated within the UNDRIP (Hogarth, 2020), namely:

> Indigenous peoples have the right to establish and control their educational systems and institutions providing education in their own languages, in a manner appropriate to their cultural methods of teaching and learning. (Article 14)

In the CANZUS countries, many measures of Indigenous socio-economic outcomes are considerably behind their equivalent non-Indigenous population (see, e.g., the global snapshot of the world's 300 million Indigenous

peoples by Anderson et al. (2016)). However, for many of the Indigenous peoples within the CANZUS countries (and elsewhere), the ability to improve their socio-economic positioning is reduced because of poor data quality. By operationalising IDSov and IDGov throughout datasets which measure socio-economic outcomes for Indigenous peoples, this will improve the data quality across the data ecosystem leading to better Indigenous outcomes.

OPERATIONALISING INDIGENOUS DATA SOVEREIGNTY (IDSOV) AND INDIGENOUS DATA GOVERNANCE (IDGOV)

Operationalising IDSov and IDGov is key to attain good Indigenous data and overcoming issues such as missing educational data (Kukutai & Taylor, 2016). In research, the notion of operationalisation is the process of defining how concepts work, whether concepts are present or absent, and how you operationalise depends on the type of work you are undertaking (Natalier, 2019). In the CANZUS countries, Indigenous peoples have done the work to define the concepts of IDSov and IDGov. While some work has begun to operationalise IDSov and IDGov in the CANZUS countries, there is still more work needed. SGD 4 ought to consider how it too can operationalise IDSov and IDGov.

Operationalising IDSov is not without challenges, and often there are barriers experienced by both Indigenous and non-Indigenous peoples when enacting IDSov (Walter, Carroll et al., 2021). These challenges occur across the entire data ecosystem, and some specific examples are:

- Tensions between Indigenous data needs and non-Indigenous data wants (Walter, 2018);
- Unaccommodating data structures (Jelfs, 2016);
- A need for greater Indigenous statistical capacity (Lovett, 2016);
- Challenges and fragility when attempting to operationalise IDSov by non-Indigenous peoples (Pool, 2016), organisations (Walter, 2016), and governments (Bishop, 2016; Jelfs, 2016).

To overcome these barriers to operationalise IDSov, a collective effort led by Indigenous peoples with support from non-Indigenous allies is required (Walter, Carroll et al., 2021). Some challenges to operationalise IDSov can be more easily overcome through education on the topic, while others need considerable resources to be specifically allocated

and larger structural changes to occur (Walter & Carroll, 2021; Walter, Carroll et al., 2021).

CONCLUSION

This chapter has argued that the concepts of IDSov and IDGov are key components for Indigenous peoples to attain good education data that reflects their worldview and educational priorities. At present, much education data on Indigenous peoples should be classified as 'missing educational data' because a large amount of the data is missing Indigenous input across each stage of the data lifecycle. To overcome this issue of missing educational data, we suggest nation states need to support calls by Indigenous peoples to operationalise the principles of IDSov and IDGov. The result will be data that appropriately reflects the Indigenous lifeworld, contributing to Indigenous peoples attaining a good education within areas that are important to them in addition to the narrower non-Indigenous measures, ultimately leading to educational outcomes that are more equitable to their non-Indigenous counterparts.

REFERENCES

ABS, 2022, Estimates of Aboriginal and Torres Strait Islander Australians, Australian Bureau of Statistics, viewed 18/12/2022. https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/estimates-aboriginal-and-torres-strait-islander-australians/latest-release.

AIATSIS, 2022, Living Languages, Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS), viewed 18/12/022. https://aiatsis.gov.au/explore/living-languages.

Andersen, C, O'Dowd, M & Gower, L, 2015, *Aboriginal Education Workers in Tasmania becoming Teachers*, More Aboriginal and Torres Strait Islander Teachers Initiative, University of Tasmania, Hobart.

Anderson, I, Robson, B, Connolly, M, Al-Yaman, F, Bjertness, E, King, A, Tynan, M, Madden, R, Bang, A, Coimbra, CEA, Jr, Pesantes, MA, Amigo, H, Andronov, S, Armien, B, Obando, DA, Axelsson, P, Bhatti, ZS, Bhutta, ZA, Bjerregaard, P, Bjertness, MB, Briceno-Leon, R, Broderstad, AR, Bustos, P, Chongsuvivatwong, V, Chu, J, Deji, Gouda, J, Harikumar, R, Htay, TT, Htet, AS, Izugbara, C, Kamaka, M, King, M, Kodavanti, MR, Lara, M, Laxmaiah, A, Lema, C, Taborda, AML, Liabsuetrakul, T, Lobanov, A, Melhus, M, Meshram, I, Miranda, JJ, Mu, TT, Nagalla, B, Nimmathota, A, Popov, AI, Poveda, AMP, Ram, F, Reich, H, Santos, RV, Sein, AA, Shekhar, C, Sherpa, LY, Skold, P, Tano, S, Tanywe, A, Ugwu, C, Ugwu, F, Vapattanawong, P, Wan, X, Welch, JR, Yang, G, Yang, Z & Yap, L, 2016, 'Indigenous and tribal

peoples' health: The Lancet-Lowitja Institute Global Collaboration', *The Lancet*, 388(10040), pp. 131–57.

Australian Curriculum, Assessment and Reporting Authority, 2020, *Measurement Framework for Schooling in Australia 2020*, viewed 18/12/2022. https://www.acara.edu.au/docs/default-source/default-document-library/measurement-framework-2020-amended-may-12-2022.pdf?sfvrsn=4ddc4c07_0.

Australian Government Department of Education, 2022, National Statement of Commitment to Transform Education, viewed 18/12/2022. https://www.education.gov.au/international-education/resources/australias-national-statement-commitment-transform-education.

Australian Government Productivity Commission, 2022, *Report on Government Services 2022: 4 School Education*, viewed 18/12/2022. https://www.pc.gov.au/ongoing/report-on-government-services/2022/child-care-education-and-training/school-education#:~:text=Nationally%20in%202021%2C%20across%20all,schools%20(86.8%20per%20cent).

Behrendt, L, 2001, 'Genocide: The distance between law and life', *Aboriginal History*, 25, pp. 132–47.

Behrendt, L, Larkin, S, Griew, R & Kelly, P, 2012, Review of higher education access and outcomes for Aboriginal and Torres Strait Islander people, Australian Government, Canberra.

Bishop, D, 2016, 'Indigenous peoples and the official statistics system in Aotearoa/New Zealand1', in T Kukutai & J Taylor (eds), *Indigenous Data Sovereignty: Towards an Agenda*, Canberra: ANU Press, pp. 291–306.

Bishop, M, Vass, G & Thompson, K, 2021, 'Decolonising schooling practices through relationality and reciprocity: Embedding local Aboriginal perspectives in the classroom', *Pedagogy, Culture & Society*, 29(2), pp. 193–211.

Bodkin-Andrews, G, Harwood, V, McMahon, S & Priestly, A, 2013, 'AIM (E) for completing school and University: Analysing the strength of the Australian Indigenous mentoring experience', in *Seeding Success in Indigenous Australian Higher Education*, Bingley, UK: Emerald Group Publishing, pp. 113–34.

Bodkin-Andrews, G, Whittaker, A, Harrison, N, Craven, R, Parker, P, Trudgett, M & Page, S, 2017, 'Exposing the patterns of statistical blindness: Centring Indigenous standpoints on student identity, motivation, and future aspirations', *Australian Journal of Education*, 61(3), 225–49.

Bodkin-Andrews, G, Foster, S, Bodkin, F, Foster, J, Andrews, G, Adams, K & Evans, R, 2021, 'Resisting the racist silence: When racism and education collide', in *Indigenous Education in Australia*, Routledge, pp. 21–37.

Brown, L, 2019, 'Indigenous young people, disadvantage and the violence of settler colonial education policy and curriculum', *Journal of Sociology*, 55(1), 54–71.

Carroll, SR, Rodriguez-Lonebear, D & Martinez, A, 2019, 'Indigenous data governance: Strategies from United States native nations', *Data Science Journal*, 18.

Carroll, SR, Garba, I, Figueroa-Rodríguez, OL, Holbrook, J, Lovett, R, Materechera, S, Parsons, M, Raseroka, K, Rodriguez-Lonebear, D & Rowe, R, 2020, 'The CARE principles for Indigenous data governance'.

Closing the Gap, 2022, *National Agreement on Closing the Gap*, COAG and NIRA, viewed 02/10/2022. https://www.closingthegap.gov.au/sites/default/files/files/national-agreement-ctg.pdf.

Davis, M, 2016, 'Data and the United Nations Declaration on the Rights of Indigenous Peoples', in T Kukutai & J Taylor (eds), *Indigenous Data Sovereignty: Towards an Agenda*, Canberra: ANU Press, pp. 25–38.

Donovan, MJ, 2015, 'Aboriginal student stories, the missing voice to guide us towards change', *The Australian Educational Researcher*, 42(5), pp. 613–25.

Dudgeon, P, Wright, M, Paradies, Y, Garvey, D & Walker, I, 2014, 'Aboriginal social, cultural and historical contexts', in P Dudgeon, H Milroy & R Walker (eds), *Working Together Aboriginal and Torres Strait Islander Mental Health and Wellbeing Principles and Practice* (pp. 3–24). Australian Government.

Espey, J, 2002, OCAP & Stewardship: A Discussion Paper for the First Nations Statistical Institute, First Nations Statistical Institute.

First Nations Information Governance Centre, 2014, *Ownership, Control, Access and Possession (OCAP™): The Path to First Nations Information Governance*, First Nations Information Governance Centre, Ottawa.

Hart, MA, 2010, 'Indigenous worldviews, knowledge, and research: The development of an Indigenous research paradigm', *Journal of Indigenous Social Development*, 1(1A).

Hart, V, Whatman, S, McLaughlin, J & Sharma-Brymer, V, 2012, 'Pre-service teachers' pedagogical relationships and experiences of embedding Indigenous Australian knowledge in teaching practicum', *Compare – a Journal of Comparative and International Education*, 42(5), pp. 703–23.

Harwood, V, McMahon, S, O'Shea, S, Bodkin-Andrews, G & Priestly, A, 2015, 'Recognising aspiration: The AIME program's effectiveness in inspiring Indigenous young people's participation in schooling and opportunities for further education and employment', *The Australian Educational Researcher*, 42(2), pp. 217–36.

Hogarth, M, 2020. 'A dream of a culturally responsive classroom', *Human Rights Defender*, 29(1), pp. 34–5.

Hudson, M., Anderson, T., Dewes, T. K., Temara, P., Whaanga, H., & Roa, T. (2017). "He Matapihi ki te Mana Raraunga" – conceptualising big data through a Māori lens. In H. Whaanga, T.T.A.G. Keegan, & M. Apperley (Eds.), *He Whare Hangarau Māori – Language, Culture & Technology.* Hamilton, New Zealand: Te Pua Wānanga ki te Ao/Faculty of Māori and Indigenous Studies, University of Waikato, pp. 64–73.

Jelfs, P, 2016, 'The Australian Bureau of Statistics' Aboriginal and Torres Strait Islander enumeration and engagement strategies: Challenges and future options', in T Kukutai & J Taylor (eds), *Indigenous Data Sovereignty: Toward an Agenda*, Canberra: ANU Press, pp. 275–90.

Kaomea, J, 2009, 'Contemplating kuleana: Reflections on the rights and responsibilities of non-Indigenous participants in programmes for Indigenous education', *AlterNative: An International Journal of Indigenous Peoples*, 5(2), pp. 78–99.

Kukutai, T & Cormack, D, 2021, '"Pushing the space": Data sovereignty and self-determination in Aotearoa NZ', in *Indigenous Data Sovereignty and Policy*, Routledge, pp. 21–35.

Kukutai, T & Taylor, J, 2016, *Indigenous Data Sovereignty: Toward an Agenda*, Camberra: ANU Press.

Lopez, JD, 2021, 'Indigenous Data Sovereignty: Quechan education data sovereignty', in *Indigenous Data Sovereignty and Policy*, Routledge, pp. 148–56.

Lovett, R, 2016, 'Aboriginal and Torres Strait Islander community wellbeing: Identified needs for statistical capacity', in T Kukutai & J Taylor (eds), *Indigenous Data Sovereignty: Toward an Agenda*, Canberra: ANU Press, pp. 213–32.

Lovett, R, Lee, V, Kukutai, T, Cormack, D, Rainie, SC & Walker, J, 2019, 'Good data practices for Indigenous data sovereignty and governance', in A Daly, SK Devitt & M Mann (eds), *Good Data*, Amsterdam: Institute of Network Cultures, pp. 26–36.

Lovett, R, Jones, R & Maher, B, 2021, 'The intersection of Indigenous Data Sovereignty and Closing the Gap policy in Australia', in M Walter, T Kukutai, SR Carroll & D Rodriguez-Lonebear (eds), *Indigenous Data Sovereignty and Policy*, London & New York: Routledge, pp. 36–50.

Maiam nayri Wingara, 2018, *Indigenous Data Sovereignty Communique: Indigenous Data Sovereignty Summit 20/06/2018*, Maiam nayri Wingara, viewed 18/12/2022. https://static1.squarespace.com/static/5b3043afb40b9d20411f3512/t/5b6c0f9a0e2e725e9cabf4a6/1533808545167/Communique%2B-%2BIndigenous%2BData%2BSovereignty%2BSummit.pdf.

Maiam nayri Wingara, 2021, *It's Time for Indigenous Data Sovereignty and Governance in Australia*, viewed 27/09/2021. https://www.maiamnayriwingara.org/.

Martin, K, 2008, Please knock before you enter: Aboriginal regulation of outsiders and the implications for researchers, Teneriffe: Post Pressed.

Martin, KL, 2017, 'Culture and identity: LSIC parents' beliefs and values and raising young Indigenous children in the twenty-first century', in M Walter, KL Martin & G Bodkin-Andrews (eds), *Indigenous Children Growing up Strong*, London: Springer, pp. 79–99.

Moodie, N, Maxwell, J & Rudolph, S, 2019, 'The impact of racism on the schooling experiences of Aboriginal and Torres Strait Islander students: A systematic review', *The Australian Educational Researcher*, 46(2), pp. 273–95.

Nakata, M, 2010, 'The cultural interface of Islander and scientific knowledge', *The Australian Journal of Indigenous Education*, 39(S1), pp. 53–7.

Natalier, K, 2019, 'Research design', in M Walter (ed.), *Social Research Methods*, South Melbourne: Oxford University Press, pp. 29–56.

National Indigenous Australians Agency, n.d. *History of Closing the Gap*, viewed 30/08/2023. https://www.closingthegap.gov.au/resources/history.

Ober, R, Bulsey, N, Pearson, N & Bartlett, C, 2021, 'Weaving Torres Strait Islander language and culture into education', in M Shay & R Oliver, *Indigenous Education in Australia*, Routledge, pp. 76–85.

Page, S, Trudgett, M, & Bodkin-Andrews, G, 2019, 'Creating a degree-focused pedagogical framework to guide Indigenous graduate attribute curriculum development', *Higher Education*, 78(1), pp. 1–15.

Paradies, Y, 2016, 'Colonisation, racism and Indigenous health'. *Journal of Population Research*, 33(1), pp. 83–96. https://doi.org/10.1007/s12546-016-9159-y

Peacock, H & Prehn, J, 2019, 'The importance of Aboriginal Education Workers for decolonising and promoting culture in primary schools: An analysis of the longitudinal study of Indigenous children (LSIC)', *The Australian Journal of Indigenous Education*, pp. 1–7.

Pool, I, 2016, 'Colonialism's and postcolonialism's fellow traveller: The collection, use and misuse of data on Indigenous people', in T Kukutai & J Taylor (eds), *Indigenous Data Sovereignty: Toward an Agenda*, Canberra: ANU Press, pp. 57–78.

Prehn, J, Peacock, H, Guerzoni, MA & Walter, M, 2020, 'Virtual tours of Country: Creating and embedding resource-appropriate Aboriginal pedagogy at Australian universities', *Journal of Applied Learning and Teaching*, 3, Special Issue 1.

Price, A, Jackson-Barrett, E, Gower, G & Herrington, J, 2017, 'Understanding the complex work of Aboriginal Education Workers in schools', *The Australian Journal of Indigenous Education*, 48(1), pp. 1–13.

Price, K, 2012, 'A brief history of Aboriginal and Torres Strait Islander education in Australia', in K Price (ed.), *Aboriginal and Torres Strait Islander Education: An Introduction for the Teaching Profession*, Port Melbourne: Cambridge University Press, pp. 1–20.

Rainie, SC, Schultz, JL, Briggs, E, Riggs, P & Palmanteer-Holder, NL, 2017, 'Data as a strategic resource: Self-determination, governance, and the data challenge for Indigenous nations in the United States', *International Indigenous Policy Journal*, 8(2).

RDA IIDSIG, 2022, *CARE Principles for Indigenous Data Governance*, RDA IIDSIG, viewed 18/12/2022. https://www.gida-global.org/care.

Rigney, LI, 1999, Internationalization of an Indigenous anticolonial cultural critique of research methodologies: A guide to Indigenist research methodology and its principles, *Wicazo sa review*, 14(2), pp. 109–21.

Ryan, L, 2012, *Tasmanian Aborigines: A History since 1803*, Sydney: Allen & Unwin.

Schnarch, B, 2004, 'Ownership, control, access, and possession (OCAP) or self-determination applied to research: A critical analysis of contemporary First Nations research and some options for First Nations communities', *International Journal of Indigenous Health*, 1(1), pp. 80–95.

SDG Tracker, 2022, Sustainable Development Goals/Quality Education, viewed 18/12/2022. https://gpseducation.oecd.org/CountryProfile?plotter=h5&primaryCountry=AUS&treshold=5&topic=EO.

Smith, DE, 2016, 'Governing data and data for governance: The everyday practice of Indigenous sovereignty', in T Kukutai & J Taylor (eds), *Indigenous Data Sovereignty: Toward an Agenda*, Canberra: ANU Press, pp. 117–38.

Suina, M & Chosa, CT, 2021, 'Growing Pueblo data sovereignty', in M Walter, T Kukutai, SR Carroll & D Rodriguez-Lonebear (eds), *Indigenous Data Sovereignty and Policy*, London & New York: Routledge, pp. 51–61.

Trudgett, M, Page, S & Coates, SK, 2022. 'Peak bodies: Indigenous representation in the Australian higher education sector', *Australian Journal of Education*, 66(1), pp. 40–56.

United Nations, 2008, *United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP)*, United Nations: Department of Economic and Social Affairs, viewed 09/09/2022. https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html.

Universities Australia, 2017, *Indigenous Strategy 2017–2020*, viewed 11/09/2022. https://www.universitiesaustralia.edu.au/wp-content/uploads/2019/06/Indigenous-Strategy-v16-1.pdf.

Walter, M, 2016, 'Data politics and Indigenous representation in Australian statistics', in T Kukutai & J Taylor (eds), *Indigenous Data Sovereignty: Toward an Agenda*, Canberra: ANU Press, pp. 79–97.

Walter, M, 2018, 'The voice of Indigenous data: Beyond the markers of disadvantage', *Griffith Review*, 60, pp. 256–63.

Walter, M & Carroll, SR, 2021, 'Indigenous data sovereignty, governance and the link to Indigenous policy', in M Walter, T Kukutai, SR Carroll & D Rodriguez-Lonebear (eds), *Indigenous Data Sovereignty and Policy*, London & New York: Routledge, pp. 1–20.

Walter, M, Carroll, SR, Kukutai, T & Rodriguez-Lonebear, D, 2021, 'Embedding systemic change – opportunities and challenges', in M Walter, T Kukutai, SR Carroll & D Rodriguez-Lonebear (eds), *Indigenous Data Sovereignty and Policy*, London & New York: Routledge, pp. 226–34.

Walter, M, Kukutai, T, Carroll, SR & Rodriguez-Lonebear, D, 2021, *Indigenous Data Sovereignty and Policy*, London & New York: Routledge.

Watson, I, 2009. Aboriginality and the violence of colonialism, *Borderlands E-journal*, 8(1), pp. 1–8.

Williamson-Kefu, M, 2022. 'Theorising knowing: A reflection on Indigenous philosophical perspectives on knowing, learning and education', *Ngoonjook: Australian First Nations' Journal*, 36, pp. 36–43.

9. Gender, missing data and SDG 4

Helen Longlands, Rosie Peppin Vaughan and Elaine Unterhalter

INTRODUCTION

One of the most striking 'success stories' in many countries over the past two decades has been the increase in girls' enrolment in and progression through primary and secondary schooling (UNDP, 2020; UNESCO, 2021, 2022b). In most low- and middle-income countries, significant progress has been achieved in gender parity, which measures the ratio of girls to boys or women to men in a given aspect of education, particularly parity in enrolment and attendance. Yet despite policy attention at national and international levels to girls' education and gender equality, including in the UN's current Sustainable Development Agenda, and the increased emphasis on and availability of data to document changes, progress towards more substantive gender equality in education beyond gender parity has been piecemeal and unclear. Significant challenges remain, such as addressing gender-based violence and the persistence of discrimination linked to racism, sexism, xenophobia and poverty (Booth, 2022; DeJaeghere, Parkes & Unterhalter, 2013; Equal Measures, 2022; Psaki et al., 2022; Unterhalter et al., 2014; Unterhalter, Robinson & Ron Balsera, 2020). Importantly, there are also problems documenting, both quantitatively and qualitatively, the nature and form of gender and intersecting inequalities in education and their significance in realising Sustainable Development Goals 4 and 5, partly because of conceptual debates and disagreements and partly because of the uneven availability of relevant data (Faul, Montjouridès & Terway, 2021).

This chapter sets out how gender equality in education appears in the Sustainable Development Goals (SDGs), the data requirements associated with the targets for SDG 4 (the education goal), and some aspects of gendered missingness associated with the current indicators. In doing

so, the chapter reports on views on inadequate or missing data that have emerged from participatory discussions and interviews held as part of the AGEE (Accountability for Gender Equality in Education) project with a range of experts working on gender equality in education in South Africa, Malawi and internationally. The chapter explores some of the 'drivers of missingness' associated with these data gaps, and outlines steps for improving data for gender equality in and through education.

GENDER AND SDG 4 MEASURES

The SDGs spotlight gender equality as a key dimension of sustainable development (Sen, 2019). SDG 5 expresses an overall vision for gender equality, while targets associated with gender equality are linked with many other goals (Equal Measures, 2022). In SDG 4, gender parity is a feature of 8 of 12 indicators.[1] Targets 4.1 – 4.3 and Target 4.5 are concerned with eliminating gender disparities in access to, participation in, and skills and knowledge development in all levels of education from pre-primary to adult and vocational education. Target 4.7 includes 'gender equality' in the list of knowledge and skills to be achieved, but there is currently no indicator for this particular dimension. For three targets – Target 4a–c – concerned with means for implementing expanded provision and improved attendance, progression and attainment, looking at infrastructure, sanitation, vocational training and teacher training and supply, there are no gender indicators. Both the narrow focus on gender parity and the lack of gender indicators for some of the targets means there are currently serious gaps in the SDG framework relating to data that can adequately measure the full complexities of gender inequalities relating to education, which has consequences not only for realising SDG 4 but the whole SDG agenda.

The targets and associated indicators for SDG 4 are, nevertheless, more comprehensive than in the Millennium Development Goals (MDGs) (2000–15), which, to some extent, demonstrates increased interest and experience in using data in educational management and governance since 2000 (Criado-Perez, 2019; Fontdevila, 2023; Scott, 2020; Sen, 2019; Smith & Benavot, 2019). The increased use of gender parity, meanwhile, partly reflects progress on collecting gender disaggregated

[1] For full list of SDG 4 targets and indicators, see: https://sdgs.un.org/goals/goal4

data linked with activism around the need for better data associated with women's rights and understandings of gender inequalities (Criado-Perez, 2019; Scott, 2020; Sen, 2019). Yet gender parity, as noted by many commentators on gender and data, does not fully capture the many complexities of gender inequalities and how these could be documented (Merry, 2016; Sen, 2019; Unterhalter, 2014). Thus, a key question is how we mobilise both the increased expertise in data use and available data to shift the focus in policy on gender equality in education beyond gender parity.

Gender parity in basic education indicators such as participation, progression and achievement is a measure widely used by educational planners, governments and campaigners. Examples include national education sector plans (ESPs), data collected and analysed by UNESCO's Institute for Statistics (UIS), and the related Global Education Monitoring (GEM) Reports published by UNESCO (e.g., UNESCO, 2018, 2020), as well as regional learning surveys, such as Programme for the Analysis of Education Systems (PASEC) and Third Regional Comparative and Explanatory Study (TERCE). Huge advantages of gender parity as a measure for education indicators are that it is conceptually clear and methodologically straightforward, it uses existing data collection systems, such as Education Management Information Systems, examination boards or household surveys, and it is easily applicable and uses comparable data across different contexts (Unterhalter, Longlands & Peppin Vaughan, 2022).

When the SDG indicator framework was confirmed, specific 'custodian agencies' were identified for each goal and assigned overall responsibility for producing related data and the means for users to access them. For most global indicators for SDG 4, the custodian agency is UIS, which has compiled and maintains a comprehensive database on SDG educational indicators, including on enrolment, attendance, completion and learning achievement in reading and mathematics, with some disaggregation depending on context.[2] UIS does not, however, collate data on a range of issues relevant to how educational experiences vary by gender, such as discrimination associated with sexual orientation, gender-based violence, social norms and values relating to gender, social and institutional gender biases, or more detailed information on intersecting inequalities, some of which are currently collected through cross-national surveys.

[2] The UIS SDG database can be explored online at: sdg4-data.uis.unesco.org

Further, because UIS has not taken a leading role to improve gender and education data (Fontdevila, 2023), this task has increasingly been taken on by other initiatives, which build on scholarly work that has used critical perspectives to explore the gender data bias (Criado-Perez, 2019; Ferrant, Fuiret & Zambrano, 2020; Scott, 2020). These include 'Women Count'[3] led by UN Women, which aims to shift how gender statistics are created, used and promoted; 'Equal Measures 2030'[4] which aims to 'connect data and evidence with advocacy and action on gender equality' and the 'Evidence for Gender and Education Resource' (EGER)[5] – an interactive database which documents research and evidence on gender and education for the global education and gender community. In most of these initiatives, however, there has been some discussion and interest in education, but limited exploration around how to improve education data beyond basic measures like enrolment and completion in formal schooling.

Overall, the SDGs represent an expanded vision of education, and gender equality in education, particularly compared to the MDGs (Wulff, 2020). In practice, however, the SDG measurement framework and the ways in which it uses existing data do not orient to a fuller engagement with the complexities of and the urgency of addressing gender inequalities associated with education.

MISSING DATA

As many critical commentaries highlight, key areas of gender associated with education, which are important for achieving SDG 4 as well as the broader Sustainable Development Agenda, are not captured in the current SDG measurement framework (Durrani & Halai, 2020; Unterhalter, 2019a). Yet little work has been done in terms of practical steps to address these gaps, either through careful review of what is missing or through systematic projects to build the required data infrastructure.

One exception is the Accountability for Gender Equality in Education (AGEE) project,[6] which is developing an innovative indicator framework for gender equality in education (see below) that can support the SDG framework and help advocate for shifting the policy focus at both national

[3] See: https://data.unwomen.org/women-count
[4] See: https://www.equalmeasures2030.org
[5] See: https://egeresource.org
[6] See: https://gendereddata.org

and international levels beyond gender parity. Through a series of critical participatory discussions and in-depth interviews, the project has, as of January 2023, consulted with over 400 representatives from government education departments, national statistical offices, civil society, academia and youth advocacy groups in South Africa and Malawi and with a wide range of international students and organisations working on gender and education issues. These discussions have provided insight into what constitutes gender inequality in education in different locations and contexts as well as what issues are more applicable across diverse contexts, what data is perceived to be key to evaluating these inequalities, and how data and indicators might be used to bring about change. In addition, a range of pressing gender issues in education for which there are insufficient or no data have been highlighted. In South Africa, for example, important data gaps noted are associated with marginalised individuals and groups; intersecting inequalities; gender-based violence and safety around schools, universities and other sites of education; discrimination in education on the basis of sexuality and/or gender identity; and opportunities, achievements and what is valued in and through education beyond foundational subjects of mathematics and literacy. Participants also noted a need for detailed, disaggregated information about government budgeting and expenditure on education, including spending on gender equality, and some documentation of what levels of cooperation exist between government departments whose work connects to social policy and practice on gender and education (such as health, and women and children's affairs). While some of these issues arise from specific concerns in South Africa, similar points have emerged in consultation on the AGEE Framework in other countries in Africa (Malawi and Sierra Leone) and in discussion with staff from international organisations working in a range of countries (Peppin Vaughan & Longlands, 2022).

Analysis of the AGEE consultations suggests missing data on gender and education not captured in the SDG framework fall into three broad categories (Peppin Vaughan & Longlands, 2022). The first category concerns data gaps within existing indicators. While participants across locations noted significant data gaps, the nature of these gaps vary greatly by region. For example, many of the gaps noted relate to the inability to disaggregate data in multiple ways: while it is usually possible to disaggregate by gender, it is typically difficult to relate this to wealth quintiles, rural/urban depending on region, or ethnicity. One example is the indicator for 4.1.1, linked to the quality of education and the skills gap

for which there is insufficient information on intersecting inequalities, as additional forms of disaggregation are not yet uniformly available.

The second category of missing data concerns data which are collected but not yet used. Exam boards, for example, collect data on girls and boys entered for and attaining in examinations, but this information is not publicly available. With the exception of predominantly high-income countries involved with international large-scale assessments in education, such as PISA and TIMSS, it is difficult to assemble data looking at how different inequalities (e.g., ethnicity, language, income, rurality and gender) intersect in academic proficiency levels. Even when such data can be generated from existing surveys such as the Demographic Health Survey/Multi Indicator Cluster Survey (DHS/MICS), there has been hardly any use of data to explore and address the intersections of gender with other areas of educational inequality (Unterhalter, Longlands & Peppin Vaughan, 2022).

The third category of missing data, noted as key to understanding and tracing gender inequalities and equality in education, are data which are not yet collected (Unterhalter, Longlands & Peppin Vaughan, 2022), such as data for the gender equality component of SDG Target 4.7, which centres on what is taught on gender equality in schools. While the target lists a broad range of knowledge and skills (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), the current indicator only captures information on global citizenship education, and education for sustainable development. Information on knowledge and skills relating specifically to gender equality would require a new form of data collection; for example, adding to an existing UNESCO questionnaire to national governments on whether it is included in education policies, curricular frameworks, teacher training and student assessments (Unterhalter, Bella & Davies, 2019).

One important data gap identified by participants in AGEE consultations from across diverse contexts is the lack of systematic measures relating to gender-based violence in and around schools. Some comparable information can be gathered from existing population-based surveys (e.g., MICS, DHS) or school-based surveys (e.g., Global School-Based Student Health Survey (GSHS), Health Behaviour in School-Aged Children Study (HBSC) and TERCE), but country coverage and survey frequency vary greatly as does data disaggregation (Delprato, Akyeampong & Dunne, 2017; Heslop, Tamez & Parkes, 2021). Recent

research by Parkes et al. (2022) reveals a distinct disconnect between qualitative and quantitative research on sexual violence in schools, which contributes to the silence around the topic; the authors argue that greater dialogue between qualitative and quantitative data would facilitate more effective policy and practice.

Participants have also frequently noted how understanding intersecting inequalities, and how they play out in local settings, is key for allocating resources appropriately and appreciating how to disaggregate when looking at national and regional averages. Other areas of missing data that were mentioned include information on girls' leadership; girls' and boys' participation in life-skills and education about climate change and the environment; access to technology in and out of schools; some of the nuance about gender issues associated with forms of financing and education budgets; gender representation in curricula and textbooks; gender and sexual identities and understandings among children and young people; and the kinds of work, roles and identities girls and boys move into after formal schooling. A further repeated theme was missing gender data regarding out-of-school children, mobile populations, refugees, people who live in illegal settlements, and other vulnerable individuals and groups who are discriminated against and marginalised and thus often missed out from official data collection processes. There is scant data on processes linked to dismantling unjust, gender inequitable structures.

The data from the AGEE discussions highlight some recurrent themes with regards to missing data: first, the ongoing lack of data is an obstacle to understanding some of the conditions of the most marginalised and discriminated against individuals and groups. Second, these gaps are reinforced by the consistent under-resourcing of national statistical offices, which are unable to work towards improvements in data despite often being aware of the gaps. Third, there is an ongoing disconnect between the planning processes and projects concerned with what data to collect, and the insights among people working on gender education issues on the ground regarding what data it is important to collect.

Some of the issues with missing gender data are organisational and institutional, but the gendered missingness of data is also linked with the politics of engaging with gender equality, women's rights, and sexuality and non-binary identities, areas which remain contested and controversial (Jolly, 2022). An overarching issue in the lack of sustained work on gender and education data is the difficulty of constructing a shared understanding of gender equality and women's rights in and through edu-

cation. Understandings of gender equality in education tend to be narrow, centring on parity of basic education indicators, such as enrolment, attendance and academic proficiency. However, this limited framing and associated measures miss a wide range of educational experiences and processes through which gender and other intersectional inequalities can have a significant effect (Unterhalter, 2019a; Unterhalter, Robinson & Ron Balsera, 2020). Moreover, within the range of feminist perspectives on gender equality and gender equality in education, there are important differences; for example, distinctions between Western, Black African, Indigenous and Islamic feminist understandings (El Omari, Hammer & Khorichide, 2021; Hokowhitu et al., 2022; Kwachou, 2023).

A more substantive approach, therefore, requires sustained discussion and a multifaceted approach, convening across diverse perspectives and considering different and additional sources of data and measures that can capture how gender and intersecting inequalities have an impact in a wider range of areas, such as the values and norms in education systems, institutional resources and processes, classroom experiences, pedagogy and curriculum, teacher training, and outcomes beyond exam results in basic subjects. The 'Call to Action on Advancing Gender Equality and Girls' and Women's Empowerment in and through Education', launched at the Transforming Education Summit at the UN in New York in September 2022, recognised the importance of engaging with a wide range of actors. But it also specifically called on governments and international agencies to invest more in data and evidence on gender and education, particularly enabling disaggregation that would allow intersecting inequalities to be tracked more effectively (United Nations, 2022). This is an important beginning, but more needs to be built around this initiative. If national governments and the international community follow this commitment, it could be the start of some significant improvements in data systems during the remainder of the SDGs and could serve to feed into further discussions of how to better link the policy and practice needed to address transformations of unjust structures and the data to document these.

In many contexts, some gender issues can be controversial and sensitive, meaning that gaining policy attention and gathering relevant data can be politically difficult – for example, identifying sexual orientation as an area of inequality. In addition, there may also be silences and shame around documenting dimensions of poverty, gender-based violence, adolescent pregnancy, female genital mutilation (FGM) and non-binary gender identities. Much clearer coordination is needed between the forms

of public policy needed to address gender injustices associated with education and the data to support this.

A further issue is that, while there has been significant mobilisation around gender data in both international, and some national women's campaigns, there is a disjuncture between these initiatives and the education sector (Peppin Vaughan, 2019). The conceptualisation of education as a sphere separate from other social relationships and areas of social policy has resulted in a separation between gender activism and education reform in many countries, and at the international level. For example, recent initiatives to improve gender statistics under SDG 5, for example, 'Women Count', and various related national initiatives (led by UN Women) have not engaged in depth with education campaigns.

Thus, missing gender data in SDG 4 are linked in with ideas, politics and processes of building a gender and education infrastructure and how data are selectively used or overlooked in public policy. We turn now to outline work we have done in the AGEE project to address some of these gaps.

ADDRESSING GAPS: THE AGEE PROJECT

In the AGEE project we have worked to delineate what further steps are needed to address data gaps on gender equality in and through education and what can be done to make better use of data that are currently available. We have integrated suggestions from the project's consultation phase (see above) in the development of a framework for understanding gender inequalities in education and progress towards gender equality and enhancing women's rights. This work has entailed developing a substantive definition of gender equality in and through education, drawing on the capability approach (see Unterhalter, Longlands & Peppin Vaughan, 2022).

The resulting AGEE Framework sets out six distinct but interconnected 'domains' that we consider (based on conceptual and empirical work) should be represented within a more holistic measurement framework for gender equality and education: Resources; Values; Opportunities; Participation in Education; Knowledge, Understanding and Skills; and Outcomes.

The Resources domain reflects the goods and services required for a gender equitable education system. Resources include, for example: funding, policies, school infrastructure, trained teachers, administrators, support workers and information. The Values domain contains informa-

tion on norms relating to gender and education, for example, provisions in constitutions, and survey data on attitudes to girls' and women's education. The Opportunities domain reflects the policy context and economic and social environment, such as laws, and national curricula. This domain includes ideas about, for example: the policy context and how this is put into practice; gendered aspects of the political, economic, geographic and cultural/social environment; and gender relations within educational institutions. The Participation domain considers gender differences in the capability to participate in education, and levels of participation and progression of girls and boys in all levels of education. Measures in this domain might include, for example: looking at differences in participation in education by socio-economic status, location, race, ethnicity. The Knowledge, Understanding and Skills domain captures information about learning across all areas of the curriculum (not just literacy and numeracy), ideally including values around rights and gender equality; and forms of learning and teaching. Finally, the Outcomes domain takes a broad definition of the results of education, for example, to include economic empowerment, political participation, and speaking out against gender-based violence (Unterhalter, Longlands & Peppin Vaughan, 2022).

By gathering information across all six AGEE domains, it is possible to provide a substantive picture of gender equality in an education system, as well as a range of forms of inequality and marginalisation. The AGEE Framework is also designed to be flexible to context, so that it is possible to get a sense of levels of inequalities in capabilities across different locations and situations.[7] In the longer term, it is envisaged that the AGEE Framework will inform data use at international, national and local levels.[8]

Work in the AGEE project between 2021 and 2023 is concerned with the global aims of the SDG measurement system and entails creating a cross-national dashboard of indicators linked to the domains of the AGEE Framework. This dashboard will allow evaluations of how suc-

[7] Flexibility and responsiveness are important because of the variability of data availability and issues across different regions and countries, for example, if completion data are hard to collect, or certain forms of violence are of particular issue in a specific context, there can be some flexibility for which measures are used.

[8] For more detail on the application of the AGEE Framework, see https://gendereddata.org

cessful initiatives and policies have been in addressing the injustices associated with the multiplicity of forms of gender inequality in education.[9] In developing this dashboard, the aim is both to construct an alternative measurement framework by drawing on existing data that have not yet been used in SDG 4, and also to lobby for further improvements in data collection in response to gaps identified.

While this cross-national dashboard is still under construction, based on analysis of the views from the expert survey, Table 9.1 shows some of the candidate indicators for each of the domains.

While there are many areas that the AGEE cross-national dashboard does not currently cover, it is a pragmatic attempt to consult on the important issues affecting gender and education, survey which data are available (and which may be available shortly, or may need to be lobbied for), and then take an appropriate selection of indicators that represent different areas of gender inequality in education that reflect a more substantive and holistic approach. Further, the process of building the dashboard has illustrated that more gender education indicators exist in some domains (e.g., Resources and Participation) than others, as it has been much harder to identify sources of data for Values relating to gender and education, Opportunities (especially those which reflect intersecting inequalities associated with ethnicity, region and poverty) and information relevant to a broader understanding of gender equality in Outcomes to education beyond exam results.

To help address the serious disconnect between public policy on gender equality in education and the ways in which data are currently used, a key aim of the AGEE project is to develop a global community of practice with actors from across different fields, to gain a more holistic perspective on gender and education data. Further, the hope is that this cross-sectoral community will be better able to lobby and advocate for improvements in data collection to address data gaps in future.

A key challenge is the level of resources that can potentially be mobilised to both collate existing data and collect new data, across the range of institutions and agencies involved in education data for SDG 4. At the national level, statistical offices are often very stretched and focused on meeting requirements of existing SDG measures and national data collection responsibilities. The exploration, collection and monitoring of new

[9] Details on the process of developing this dashboard can be found on the project website, https://gendereddata.org

Table 9.1 Candidate indicators for the AGEE cross-national dashboard, and data which are not currently able to be included

Domain	Examples of candidate indicators that meet the criteria for inclusion in the AGEE cross-national dashboard and the number of countries for which data are available	Examples of data which are available but require intensive work to compile from a range of sources	Examples of data which are not currently collected in a form that meets the AGEE criteria
Resources	Proportion of total government spending on education (SDG indicator 1.a.2; 156 countries with at least one data point for the period 2010–19) Proportion of population with access to a mobile phone network (SDG indicator 9.c.1; data for this indicator exist for more than 160 economies)	Household expenditures on education as a proportion of average household income	Proportion of budget of education which is gender responsive/ transformative
Values	Women and men over 15 guaranteed full and equal access to sexual and reproductive health care, information and education (SDG indicator 5.6.2; 153 countries have completed or partial data)	[No indicators currently identified]	Attitudes on whether boys are more suited to STEM than girls

Domain	Examples of candidate indicators that meet the criteria for inclusion in the AGEE cross-national dashboard and the number of countries for which data are available	Examples of data which are available but require intensive work to compile from a range of sources	Examples of data which are not currently collected in a form that meets the AGEE criteria
Opportunities	Proportion of children aged 5–17 years engaged in paid child labour (SDG indicator 8.7.1; nationally representative and comparable data are currently available for around 100 low- and middle-income countries)	Percentage distribution of households and de jure population by time to obtain drinking water	Positions/management positions of female teachers
Participation in Education	Adjusted GPI of GER in tertiary education (UIS/ SDG indicator 4.3.2; widely available from administrative data)	Proportion of girls in three lowest socio-economic status (SES) quintiles whose mother has completed primary school, and proportion who have completed secondary school (calculate from DHS)	Measures bringing together conflict status with refugees and participation in learning

Domain	Examples of candidate indicators that meet the criteria for inclusion in the AGEE cross-national dashboard and the number of countries for which data are available	Examples of data which are available but require intensive work to compile from a range of sources	Examples of data which are not currently collected in a form that meets the AGEE criteria
Knowledge, Understanding and Skills	Gender Parity Index (GPI) of students enrolled in STEM courses (UIS; varied availability) GPI of students accessing ICT skills training (SDG indicator 4.4.1; as of 2020, 91 economies have ever reported ICT skills data since 2005).	Women enrolled in research degrees as a proportion of women with undergraduate degrees	Learning outcomes relating to gender equality
Outcomes	Proportion of women in positions in national and local institutions, including (a) the legislatures; (b) the public service; and (c) the judiciary, compared to national distributions, by sex, age, disability and population group (SDG indicator 16.7.1; availability varies by component but often available for 193 countries) Proportion of time spent by women and men on unpaid domestic and care work (SDG indicator 5.4.1; 90 countries with data between 2000 and 2020)	Access to essential health services for women (taking account of SES, location and level of need – may need calculation from DHS)	Speaking out about gender-based violence (GBV) (e.g., presence of girls' and women's voices on aspects of school-related GBV) Women's leisure time/ GPI leisure time

Source: UN Stats SDG Metadata portal (2022) [https://unstats.un.org/sdgs/dataportal/SDMXMetadataPage, accessed December 2022] and UIS database.

data sources may thus be a step too far unless more investment is made in national statistical offices and building education departments' capacity. But the lack of detailed data on gender and other intersecting inequalities both reflects and limits the opportunities to integrate a more substantive gender perspective in national government and statistical institutions (Fukuda-Parr, 2019; Jerven, 2019).

Similarly, international agencies often have limited capacity to undertake the harder task of producing additional gender data. Projects from aid agencies, even those that are large and have a focus on girls' education, may be irregular or limited to narrow monitoring, evaluation and learning (MEL) frameworks, rather than working up more nuanced metrics, because of wider accountability structures they are part of. Likewise, civil society organisations, especially smaller, grassroots entities which have the relevant experience of the most pressing gender issues in education in particular contexts are likely to have limited resources to either collect new data or lobby for improved measurement frameworks.

Compounding this issue, particularly since 2015, is that many contemporary crises relating to health, climate, conflict and political and economic instability have brought substantial disruption to even basic data collection processes, while at the same time creating the need for additional measures to adequately capture the gendered effects of these events. Conflict and climate-related emergencies in specific contexts, for example, have severely affected the capacity of national statistical agencies to gather routine statistics (Fukuda-Parr, 2019; Jerven, 2019). Emerging research has demonstrated the gendered effects of climate injustices on education, which are not currently measured (Pankhurst, 2022). The Covid-19 pandemic has had a global impact on the collection of education data, while simultaneously creating an urgent need to gather new information to understand the gendered effects on children's learning and wellbeing (UNESCO, 2022a).

CONCLUSION

This chapter has outlined the ways in which gender does and does not appear in SDG 4, highlighting key aspects of gender and intersecting inequalities in and through education on which data are missing or inadequate. These data gaps hinder processes to better understand, analyse, monitor and address the many and complex gender injustices associated with education and various connected areas of social policy. These gaps, therefore, have consequences for the wellbeing of individuals, social

relations and societies, and the achievement of the overall orientation of SDG 4, SDG 5 and the broader sustainable development agenda. There is extensive work to be done at cross-national, national and local levels. Particular forms of coordination are required between governments, organisations, institutions and individuals that do not validate harm, violence or reinscribe inequalities, but rather, aim to pool resources and knowledge to dismantle deeply entrenched forms of injustice. Positioning gender as a key element is a vital step to progress.

REFERENCES

Booth, R. (2022). Emerging Issues in East Africa for Girls' Education in East Africa. *K4D Emerging Issues Report 48*. Brighton, UK: Institute of Development Studies (IDS).

Criado-Perez, C. (2019). I*nvisible Women: Exposing Data Bias in a World Designed for Men*. London: Vintage.

DeJaeghere, J., Parkes, J., & Unterhalter, E. (2013). Gender Justice and Education: Linking Theory, Policy and Practice. *International Journal of Educational Development, 33*(6). doi:10.1016/j.ijedudev.2013.05.006

Delprato, M., Akyeampong, K. & Dunne, M. (2017). The Impact of Bullying on Students' Learning in Latin America: A Matching Approach for 15 Countries, *International Journal of Educational Development, 52*, 37–57

Durrani, N., & Halai, A. (2020). Gender Equality, Education, and Development: Tensions between Global, National, and Local Policy Discourses in Postcolonial Contexts. In A. Wulff (ed.), *Grading Goal Four: Tensions, Threats, and Opportunities in the Sustainable Development Goal on Quality Education*. Leiden: Brill. 65-95

El Omari, D., Hammer, J., & Mouhanad, K. (eds) (2021). *Muslim Women and Gender Justice: Concepts, Sources, and Histories.* Abingdon, UK: Routledge.

Equal Measures. (2022). 'Back to Normal' Is Not Enough: the 2022 SDG Gender Index Woking: Equal Measures 2030.

Faul, M., Montjouridès, P., & Terway, A. (2021). Education and the New Inequality Divides. *Global Challenges (NORRAG),* 9.

Ferrant, G., Fuiret, L., & Zambrano, E. (2020). The Social Institutions and Gender Index (SIGI) 2019: A Revised Framework for Better Advocacy. *OECD Development Centre Working Papers*, No. 342, OECD Publishing, Paris, https://doi.org/10.1787/022d5e7b-en.

Fontdevila, C. (2023). The Politics of Good Enough Data. Developments, Dilemmas and Deadlocks in the Production of Global Learning Metrics. *International Journal of Educational Development, 96*. doi:10.1016/j.ijedudev.2022.102684

Fukuda-Parr, S. (2019). Keeping out Extreme Inequality from the SDG Agenda – the Politics of Indicators, *Global Policy, 10*(S1), 61–9. doi:10.1111/1758-5899.12602

Heslop, J., Tamez, L.Q., & Parkes, J. (2021). Appraisal of Data Available on Violence Against Children in and around Schools (VACS). *UCL Institute of Education Working Paper*.

Hokowhitu, B., Moreton-Robinson, A., Tuhiwai-Smith, L., Andersen, C., & Larkin, S. (eds), *Routledge Handbook of Critical Indigenous Studies.* Abingdon, UK: Routledge.

Jerven, M. (2019). The Problems of Economic Data in Africa. In W.R. Thompson (ed.), *Oxford Research Encyclopedia of Politics.* Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.748

Jolly, S. (2022). Is Development Work Still So Straight? Heteronormativity in the Development Sector over a Decade on. *Development in Practice.* doi:10.1080/09614524.2022.2115012

Kwachou, M. (2023). In Response to Acker: Black and African Feminist Theories on Gender and Education. *Comparative Education*, *59*(2), 169-192. DOI: 10.1080/03050068.2022.2147634

Merry, S.E. (2016). *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*. Chicago, IL: University of Chicago Press.

Pankhurst, C. (2022). Girls' Education and Climate Change: A Critical Review of the Literature. *Accountability for Gender Equality in Education (AGEE) Working Paper.* London: Centre for Education and International Development, UCL.

Parkes, J., Bhatia, A., Datzberger, S., Nagawa, R., Naker, D., & Devries, K. (2022). Addressing Silences in Research on Girls' Experiences of Teacher Sexual Violence: Insights from Uganda. *Comparative Education*, 1–21. doi:10.1080/03050068.2022.2133861

Peppin Vaughan, R. (2019). Global Campaigns for Girls' and Women's Education, 2000–2017: Insights from Transnational Social Movement Theory. *Comparative Education, 55*(4), 494–516. doi:10.1080/03050068.2019.1657701

Peppin Vaughan, R., & Longlands, H. (2022). A Technology of Global Governance or the Path to Gender Equality? Reflections on The Role of Indicators and Targets for Girls' Education. *Comparative Education*, 1–23. doi:10.1080/03050068.2022.2145774

Psaki, S., Haberland, N., Mensch, B., Woyczynski, L., & Chuang, E. (2022). Policies and Interventions to Remove Gender-Related Barriers to Girls' School Participation and Learning in Low- and Middle-Income Countries: A Systematic Review of the Evidence. *Campbell Systematic Reviews, 18*(1), e1207.

Scott, L. (2020). *The Double X Economy: The Epic Potential of Empowering Women* (Export edn). London: Faber & Faber.

Sen, G. (2019). Gender Equality and Women's Empowerment: Feminist Mobilization for the SDGs. *Global Policy, 10*, 28–38.

Smith, W.C., & Benavot, A. (2019). Improving Accountability in Education: The Importance of Structured Democratic Voice. *Asia Pacific Education Review, 20*(2), 193–205. doi:10.1007/s12564-019-09599-9

UNDP. (2020). *Human Development Report 2020. The Next Frontier: Human Development and the Anthropocene*. New York: UNDP.
UNESCO. (2018). *Global Education Monitoring Report: Accountability in Education: Meeting Our Commitments.* Paris: UNESCO.
UNESCO. (2020). *Global Education Monitoring Report: Inclusion and Education: All Means All.* Paris: UNESCO.
UNESCO. (2021). *Global Education Monitoring Report 2021/2: Non-State Actors in Education: Who Chooses? Who Loses?* Paris: UNESCO.
UNESCO. (2022a). *Evidence on the Gendered Impacts of Extended School Closures: A Systematic Review*. Paris: UNESCO.
UNESCO. (2022b). *Global Education Monitoring Report – Gender Report: Deepening the Debate on Those Still Left behind*. Paris: UNESCO.
United Nations. (2022). Call to Action on Advancing Gender Equality and Girls' and Women's Empowerment in and through Education. New York: United Nations Transforming Education Summit, September.
Unterhalter, E. (2014). Measuring Education for the Millennium Development Goals: Reflections on Targets, Indicators, and a Post-2015 Framework. *Journal of Human Development and Capabilities, 15*(2–3), 176–87. doi:10.1080/19452829.2014.880673
Unterhalter, E. (2019a). The Many Meanings of Quality Education: Politics of Targets and Indicators in SDG4. *Global Policy, 10*(S1), 39–51. doi:10.1111/1758-5899.12591
Unterhalter, E., North, A., Arnot, M. et al. (2014). *Interventions to Enhance Girls' Education and Gender Equality: A Rigorous Review of Literature*. London: Department for International Development.
Unterhalter, E., Bella, N., & Davies, J. (2019). *Making gender equality a reality in education: what will it take?* Washington, DC: Global Partnership for Education. Blog post.
Unterhalter, E., Robinson, L., & Ron Balsera, M. (2020). The Politics, Policies and Practices of Intersectionality: Making Gender Equality Inclusive in and through Education. *Background paper prepared for the Global Education Monitoring Report, 2020 Gender Report: A New Generation: 25 Years of Efforts for Gender Equality in Education.*Paris: UNESCO.
Unterhalter, E., Longlands, H., & Peppin Vaughan, R. (2022). Gender and Intersecting Inequalities in Education: Reflections on a Framework for Measurement. *Journal of Human Development and Capabilities*, 1–30. doi:10.1080/19452829.2022.2090523
Walker, M. (2014). Educational Transformation and Gender Justice. In F. Comim & M. Nussbaum (eds), *Capabilities, Gender, Equality.* Cambridge: Cambridge University Press.
Wulff, A. (ed.) (2020). *Grading Goal Four: Tensions, Threats, and Opportunities in the Sustainable Development Goal on Quality Education*. Leiden: Brill.

10. Conclusions on missing education data and the SDG 4 data regime

Marcos Delprato and Daniel Shephard

INTRODUCTION

With only seven years left to reach the SDGs' agenda deadline, this book has shown that the achievement of SDG 4 remains uncertain in part because of missing data within its monitoring architecture. Such data gaps often affect contexts and groups that are furthest left behind in terms of reaching the SDG 4 goals due to compounded disadvantages operating at different levels (UNESCO, 2020). The book has also highlighted the misalignment that exists at times between the global SDG 4 data regime and the priorities at national and regional levels. The book's different chapters have presented evidence that missing data and inadequate data use can present significant challenges for international, national, and local education stakeholders seeking to achieve SDG 4 targets (Arias Ortiz et al., 2019; Subosa & West, 2018). The chapters have presented examples of education data gaps that are caused by non-existent data, unused data, data that lack comparability, and data that are misaligned with the needs of educational stakeholders. The leading message from the book is that missing data hamper a robust data picture on what underlies educational inequalities and, thus, it becomes a crucial obstacle towards fulfilling SDG 4.

Embedded inequalities are crucial barriers for resource-constrained education systems, especially the so-called least-developed countries which are furthest away from achieving the SDG 4 targets (UNDESA, 2022). Unfortunately, these same contexts also experience the largest gaps in terms of data availability, quality, comparability, and use that could inform targeted education policies (UNDESA, 2022; World Bank, 2021). Yet, the opportunity to improve the identification of those furthest marginalized – and the mechanisms leading to such marginalization

– can help accelerate the achievement of SDG 4 (Grek, 2022). This requires a critical appraisal of the elements of the SDG 4 monitoring architecture and better incorporation and utilization of existing datasets and their possible disaggregation.

There is also the deeper question of whether SDG 4 and its data regime only represent a global monitoring engine for accountability, or if they can support the achievement of the SDG 4 goal by informing changes in policy and practice, for example, through facilitating cross-national learning and amplifying civil society advocacy. Critiques of the SDG 4 targets and data regime range from the effectiveness of global goals, to critiques of quantitative governance that can exclude through categorization, and epistemological critiques that such global education development goals and data regimes marginalize local epistemologies (Hoppers, 2015; Klees, 2017; Normand, 2020; Ntihirageza & Ibrahima, 2021; Olusanya et al., 2021). This, in turn, is connected to the challenges highlighted by the chapters on the relative importance of SDG 4 data alignment for educational policies at the national, local, or community levels. In other words, one question tackled throughout the book is about data production vis-à-vis data purposes and, ultimately, data use.

Similarly, the chapters raise concerns about SDG 4 indicators not fitting certain contexts, not being timely, or not appropriately measuring ingrained sources of educational inequalities, such as Indigeneity (Chapter 8) and gender (Chapter 9), where in both cases a participatory bottom-up approach and more careful definition of indicators could be more successful (Burford, Tamas, & Harder, 2016). For example, Chapter 9 argued that "the SDG measurement framework and the ways in which it uses existing data do not orient to a fuller engagement with the complexities of and the urgency of addressing gender inequalities associated with education" (p. 141). Also, as Chapter 7 highlighted, the lack of timely, reliable, and comparable data on IDPs' education which "hamper the design and implementation of tailored policies and programmes" (p. 116).

The remainder of this final chapter includes a summary of the missing monitoring data framework proposed in Chapter 2 and its connection with the different regional and thematic chapters. In addition, we draw out some of the shared themes across the book in terms of missing data and gaps in the use of data to inform policies and programmes to achieve quality education for all.

MISSING FRAMEWORK FOR MONITORING OF SDG 4

Motivated by accounting for the multidimensionality process of entrenched education inequalities, Chapter 2 of this book introduced a heuristic framework relying on current available information from learning and household surveys used to monitor many of the SGD 4 targets. This framework is an attempt to represent layered sources of inequalities more granularly within an ecological framework, starting at the child/youth level and then moving-up to their households, schools, and communities, and, importantly, using drivers and their intersections, where these intersections are constructed by concentric layers of disadvantages shaping a given education indicator.

This book investigates three types of missing categories in relationship to how SDG 4 indicators are monitored in UNESCO platforms: (1) data that are missing because of reporting (MR) in which data exist but are not reported, including unreported disaggregation or entire unreported indicators; (2) data that are missing because of the absence (MA) of data collection; and (3) data that are missing because of the monitoring framework (MF), where key drivers of SDG 4 targets exist in underlying datasets, especially household surveys and learning assessments, but are not part of the SDG 4 monitoring framework. The book's chapters provide instances of all three types of missingness.

Chapter 2 of the book showed the link of the heuristic framework to assess the extent of the problem of missing data in the monitoring of SDG 4, and it argued that overlooking dimensions and their combinations among drivers used for monitoring can lead to wider education deprivation. It further demonstrated that the exclusion of combinations is more problematic in the case of learning indicators compared to completion (or access) indicators. For instance, for learning outcomes (target 4.1.1), UIS data do not use available two dimensions (2D) or three dimensions (3D) of disadvantages affecting learning. The chapter also drew our attention to layers of inequality put forward by the heuristic framework that provide a chance to address marginalization by incorporating them into monitoring systems. Vitally, this new framework reminded us of the importance of intersecting inequalities to achieve SDG 4 and its intersection with numerous SDGs, such as SDG 5 on gender inequality.

MISSING INDICATORS AND EDUCATION STAGES ACROSS REGIONS

The regional book chapters (Chapters 3 to 6) offered evidence on the wide array of challenges regarding the monitoring of SDG 4 in Africa, Asia, the Arab States, and Latin American and the Caribbean. These challenges range, for example, from poor data recording coupled with a lack of financial resources and technical expertise (Africa); weak statistical systems coupled with the challenges presented by monitoring and acting on educational data in contexts affected by conflict and large-scale displacement (Arab States); invisible children, those attending school in conflict-affected regions, and non-government schools (Asia); and data gaps when it comes to learners on the move and those with disabilities alongside challenges of poor data integration across countries, sectors, and systems (Latin America and the Caribbean).

All regions have the common data problem of limited comparability and a lack of regional benchmarks, although there have been recent efforts to establish these. In addition, all four regions face a scarcity of data disaggregation to capture education (in)equity. Given this scenario, a way forward is to reflect on common threads between chapters embedded in the three types of missing data. Data that are missing because of reporting (MR) are associated with comparability and the capabilities of data processing and analysis, including financial resources and expertise. Data that are missing because of the framework (MF) relate to the lack of data use when data are available. This lack of use can be due to insufficient integration (e.g., with data systems outside of education) or available data that does not precisely capture intersecting drivers of regional educational inequalities. Finally, data that are missing due to weak statistical systems and the absence of benchmarks and standards relate to absence of data (MA).

The MR type was identified as a key concern in Africa where "fewer than a quarter of countries in Africa have provided the UIS with SDG 4.1 data since 2014" (p. 86) and, even if reported, there is a lack of "disaggregation of learners by different types of vulnerability, for both in- and out-of-school populations" (p. 88). Worryingly, too, regarding SDG 4 indicators related to learning achievement, 17 out of the 54 (nearly a third) African countries have not carried out learning surveys, and only 15 countries have undertaken national assessments at secondary level resulting in data missing due to absence (MA). Insufficient funding

of statistical offices is a prominent driver of these gaps that needs to be addressed – for education data and more broadly for national statistics offices in line with SDG 17.19. Equally, the Arab States are severely affected by data missing due to absence (MA) because less than half of the countries are collecting the needed data to produce global indicators and 43 per cent of disaggregated indicators are unavailable (p. 70). This is explained by a lack of statistical capacity – driven both by poor training in data collection and use alongside a lack of funding which deepens the lack of human capacity; continued weak data infrastructure; and disrupted data collection due to conflict, natural disasters, and large-scale displacement. The Asian chapter showed a lack of reporting at the start of the SDG 4 agenda with only 44 per cent of the global indicators collected with large gaps continuing to this day. The chapter also demonstrated continued gaps in terms of disaggregation and parity indices. For example, Chapter 4 emphasized that for five countries in Asia and the Pacific, "no country had a complete set of the 12 analysed education indicators with sex disaggregation and only 55 per cent of the education indicators were conforming and sex-disaggregated" (p. 57). Latin America and the Caribbean, even though the region has a greater monitoring capacity according to some metrics (World Bank, 2021), also faced a challenge of MA because of the inclusion of less traditional/new indicators that extend beyond basic education: "the enrolment ratios in pre-primary or tertiary education, youth enrolled in Technical and Vocational Education and Training (TVET), or the literacy rate of youth and adults" (p. 38).

Throughout regions leading missing groups were early childhood learners, youth, and a broad range of marginalized groups which depend on each region's context. Even in the region with the highest data production and highest mean years of compulsory pre-primary education guaranteed in legal frameworks (i.e., LAC), there is a lack of information on early childhood education due to the multisectoral nature of early childhood education (ECE) services, which results in data scattered across government and non-governmental entities that often lie outside of traditional EMIS systems focused on government-run basic education levels. In the Arab States, data collected for this target using household surveys are scarce with enormous between-country variation. In Africa, too, access to ECE varies extensively across the continent, and this affects monitoring of these early learners.

Youth and adults in education and training are also a group facing systematic data gaps across regions because of the complexities of

gathering information from non-formal education, tertiary education, and technical and vocational training (TVET) programmes outside of basic education and often offered outside the public education sector. Yet this is a priority group of learners for many regions. For example, in Africa, where data are available, "58 per cent of youth of secondary school age not enrolled in secondary school are likely to be in some form of employment, apprenticeship, or non-formal training" (p. 90). Meanwhile, across the Arab States around 18 per cent of the youth are not in employment, education, or training (NEET). In most regions, TVET is highly fragmented, and most countries have a large number of institutions with different structures of ownership and control, as discussed in Chapter 4 on Asia. This pattern of institutional fragmentation emerges in the Latin American and the Caribbean region as well and requires setting up new frameworks for the inclusion of the array of learning programmes in information systems beyond traditional educational institutions in order to include formal, informal, and non-formal education. The complexity of TVET invariably has led to its underrepresentation in formal educational settings and, therefore, to a significant gap on missing data due to absence and reporting.

MISSING MARGINALIZED GROUPS

Many marginalized groups suffer from the perpetuation of educational inequalities while also being poorly represented in disaggregated data (Olusanya et al., 2021; UNDESA, 2022; UNESCO, 2020). These are the groups who are furthest from realizing equal access to quality education for sustainable development. However, there are important regional, national, and local variations on which groups are most marginalized in education. In some cases, this marginalization is a product of long-standing historical inequalities and epistemic violence, as is the case with Indigenous peoples in both the Global North and the Global South. In other cases, groups are marginalized due to contemporary shocks, such as wars, natural disasters, or economic crises. Finally, some groups are marginalized due to the structure of the global and national education architecture which has traditionally prioritized basic education for children within government-run schools.

Throughout both the regional (Chapters 3–6) and thematic chapters (Chapters 7–9) several common themes of marginalized groups emerge. First, there are historically excluded groups who have often been excluded from political power structures, notably: those who

are differently-abled, female learners, non-citizens, and Indigenous peoples. Second, there are those who have been uprooted by shocks that have rippled through education systems, notably: learners who live in conflict-affected and fragile settings, have been forcibly displaced (either internally or across international boundaries), or have been affected by economic upheavals – including forced economic displacement. Finally, there are those who have been traditionally de-prioritized in international and national architectures that pre-date the SDGs, notably: early learners before primary schooling, youth and adult learners, and learners outside of the public school system as noted above.

In some cases, chapters highlighted that the measurement of marginalized groups, once incorporated and introduced into monitoring frameworks, have presented further difficulties for data capture, governance, and use. For instance, most approaches to collecting data about students with disabilities do not use a social approach and thus cannot inform the identification of the specific needs among students with difference disabilities. This translates into difficulties in the design and implementation of inclusive and relevant educational practices. Similar challenges in data production and use face internally displaced persons (IDPs). There are technical difficulties in collecting education data among IDPs, including the challenge of creating sampling frames that enable representative data collection, the fact that data are quickly outdated due to the groups' relative mobility and desire (at times) to avoid identification, and the reality that public schools and government data systems may not be functional in areas affected by conflict, disaster, and displacement. These difficulties, coupled with the political challenges of identifying and addressing the needs of IDPs in many contexts, create substantial barriers for reporting on and providing quality education for the tens of millions of learners currently living in displacement. In the case of Indigenous peoples, Chapter 8 provided a powerful argument for the need to ensure that data governance is in the hands of the communities who are affected by that data – only then will data be useful for improving education for all.

MISSING EDUCATION DATA USES

When it comes to data use, the book demonstrates many common challenges related to:

- Inadequate data alignment across local, national, regional, and global levels;

- Misalignment of local/national data needs for education policymaking with the SDG 4 monitoring architecture;
- Uneven data capacity at three levels: financing of national and educational statistics' capacities, hiring and retaining skill staff, and strengthening civil societies' capacity to access, analyze, and use data;
- Path dependencies in education data systems and the resulting difficulties of reforming them; and
- Weak systems in support of data use that can bridge SDG 4 indicators and goals.

One reason for this is that when the new SDG architecture was created, educational data production systems were neither prepared to collect the data, much less to process and use it effectively. Indeed changes to the indicator framework have continued throughout the first half of the SDG 4 agenda (Hereward, 2021; Unterhalter, 2019). Therefore, systems of data production and use in education have required substantial investment – however, this investment has been lacking both in terms of global support and in terms of national investments as noted across the regional chapters and in related publications on SDG 4 data systems (UNDESA, 2022; World Bank, 2021). Finally, there is a weak culture of data sharing and use at all levels – global, regional, national, and subnational. This results in disconnected sources of data that would need to be combined to provide robust policy inferences to address continued educational shortcomings and inequalities.

Furthermore, education data collection, analysis, dissemination, and use respond to different aims such as planning, policies design, monitoring, and evaluation all taking place at different structures (central, local, etc.) and locations of education systems (Buckner, Shephard, & Smiley, 2022; Burford, Tamas, & Harder, 2016; Grek, 2020; World Bank, 2021). This makes the unification of the varied data sources and their alignment with the needs of data users quite challenging, with education stakeholders, knowledge brokers, and government officials attempting to amalgamate sources based on their needs, the data availability, and their capacity.

In order to develop a robust picture of the educational needs of any system it is necessary to draw from the Education Management Information Systems (EMIS) alongside household surveys and assessment data. This process can be substantially facilitated through the use of individual-level data – although this raises important data-privacy

concerns (UNICEF, 2021). For instance, in Latin America, the use of student IDs (i.e., the introduction of individual student-level data) has increasingly enabled the linkage of EMIS with other databases within and outside of the Ministries of Education to facilitate data use and policy action. This has the potential of offering a virtuous circle of education data production and use. Yet, as argued in Chapter 3, there is a long way to go in this respect as data dissemination within the production cycle is a low priority among the different agencies in charge of educational data, therefore hampering the design of effective and responsive educational policies.

Household surveys produce indicators on important additional population-level data on education (e.g., participation and completion, educational attainments of adults, and literacy rates) and, by doing so, they are a useful resource to compensate for some missing data that go beyond the traditional EMIS boundaries (e.g., surveys include out-of-school populations). Surveys also have the added value of providing measures of equity in education (and its drivers) as they include information on gender, family wealth, location, etc. However, household surveys also face challenges in the identification of vulnerable population – especially when the survey is national and the population is relatively small and/or not included in sampling frames. In addition, surveys face the added complication of being irregular and expensive to conduct. Household surveys often lack representative subsamples of undocumented and displaced people, for example. Finally, national surveys with data relevant for education are only as reliable as their underlying sampling frames, and many countries lack regular and reliable censuses and vital statistics (World Bank, 2021).

Learning assessments are used to calculate indicators for learning outcomes (mostly, maths and reading achievement) for SDG 4.1. They include international learning assessments (such as PIRLS, PISA, TIMSS); regional assessments gathering representative regional-level data using a common approach (e.g., ERCE in Latin America, PASEC for the Francophone education systems in Africa); and national assessments which can offer better insights on learning drivers being more aligned to the realities of national education systems, though they are costly and require alignment with global frameworks to enable comparability and reporting on SDG 4 (UNESCO, 2019). Unfortunately, international, regional, and national assessments are not widely available for all countries; regional and international assessments are infrequent and their sample frames often only allow the construction of national statis-

tics – not the subnational data needed by many education policymakers. In addition, national assessments tend to be focused at different grades of primary or secondary and are not comparable between countries – despite recent efforts to overcome this (UNESCO, 2019).

Regional assessments are arguably a better tool for comparison, as they are aligned to regional realities and enable cross-country comparisons. However, they are expensive and require strong technical expertise and have other sample-based constraints. Learning surveys, overall, have the extra limitation of gathering limited competencies among the youth and adult population, and are often missing other learning areas such as "transferable skills, global citizenship education or education for sustainability, all core issues for SDG 4 monitoring" (p. 42). However, regional surveys such as ERCE have increased their scope, for example, piloting modules for measuring soft skills, global citizenship education, and education for sustainable development. It should be emphasized, nonetheless, that there is highly variable country coverage of these types of assessments per region. As mentioned above, in the case of Africa, 31 per cent of countries have not conducted any learning assessment since 2014 and, for the Arab region's least developed countries, there are no data available for reading and mathematics proficiency at the end of lower secondary.

SDG 4, EDUCATION DATA, AND MOVING FORWARD

Cross-country comparisons of educational performance and progress on SDG 4 indicators are a significant and costly endeavour. There are only seven years left for the SDGs' agenda deadline, and the accomplishment of SDG 4 remains uncertain. Education systems have unique problems and barriers with inequality manifesting differently according to specificities of these systems and countries' realities. However, the homogenization of SDG 4 indicators is aimed at allowing cross-country comparison even though deep educational inequalities are often – as shown in this book's different regional chapters – entrenched mechanisms with local dimensions.

Meaningful data use for impactful educational policies is a product of a combination of data sources, education stakeholders, their purposes, and the context. SDG 4 indicators, being homogeneous, act as umbrella for a supra-national data monitoring regime of education access, quality, and (in)equality, highlighting common barriers facing all countries. Yet,

even for this generic and global monitoring framework of education, the issue of missing data can sever the required link between data sources and data purposes. Missing data will operate in a different way across countries, often leaving behind those groups experiencing wider marginalization and compounded inequalities.

A fundamental question remains of whether the SDG 4 education goal and its data regime can be thought of as more than simply a global monitoring apparatus and what modifications are needed for the data regime to more effectively support the achievement of the goals of SDG 4. Is it possible for the global data regime to have an impact at the national and subnational levels where education policies are crafted and implemented? Although international educational data alignment is important, the various bottlenecks linked to SDG 4 education indicators (whether comparability, lack of capacity, costs of surveys, etc.) may suggest the need to prioritize a "bottom-up" approach to global education data. First, greater efforts could be targeted towards supporting stakeholders' capacity for within-country data comparability and quality regarding local sources of education inequalities across subregions and time. Then, only secondarily, should we support cross-country alignment with SDG 4 indicators that, ultimately, may not be as useful for country-level policymaking.

Ultimately, it will be the effective use of data – not merely its existence – that supports the achievement of SDG 4, and this requires investing in data production, data access, and data use in a way that supports those who need the data to improve their education systems. This must include all education stakeholders including policymakers, practitioners, communities, and learners. We hope this book has helped move the conversation in that direction and will help add momentum to creating a more effective global education data regime that can help drive the achievement of quality education for all.

REFERENCES

Arias Ortiz, E., Eusebio, J., Pérez Alfaro, M., Vásquez, M., & Zoido, P. (2019). *From paper to the cloud. Guiding the digital transformation of Education Management and Information Systems*. Washington, DC: Inter-American Development Bank (IADB).

Buckner, E., Shephard, D.D., & Smiley, A. (2022). Beyond numbers: The use and usefulness of data for education in emergencies. *Journal on Education in Emergencies, 8*(1). doi:https://doi.org/10.33682/tgfd-m9eg

Burford, G., Tamas, P., & Harder, M.K. (2016). Can we improve indicator design for complex Sustainable Development Goals? A comparison of a values-based and conventional approach. *Sustainability, 8*(9), 861. doi:https://doi.org/10.3390/su8090861

Grek, S. (2020). Prophets, saviours and saints: Symbolic governance and the rise of a transnational metrological field. *International Review of Education, 66*(2), 139–66. doi:10.1007/s11159-020-09844-z

Grek, S. (2022). The education Sustainable Development Goal and the generative power of failing metrics. *Policy and Society, 41*(4), 445–57. doi:10.1093/polsoc/puac020

Hereward, M. (2021). Setting the record straight on measuring SDG 4.2.1. *The Lancet Global Health, 9*(7), e910–e911. doi:10.1016/s2214-109x(21)00229-1

Hoppers, C.A.O. (2015). Think Piece: Cognitive justice and integration without duress. The future of development education – perspectives from the South. *International Journal of Development Education and Global Learning, 7*(2), 89–106.

Klees, S.J. (2017). Will we achieve Education for All and the Education Sustainable Development Goal? *Comparative Education Review, 61*(2), 425–40. doi:10.1086/691193

Normand, R. (2020). The politics of metrics in education: A contribution to the history of the present. In G. Fan & T.S. Popkewitz (eds), *Handbook of education policy studies* (pp. 345–60). Singapore: Springer.

Ntihirageza, J., & Ibrahima, A. (2021). Education and epistemicide in Africa: Towards an Ubuntu-based comprehensive model of education. In R. Tshifhumulo & T.J. Makhanikhe (eds), *Handbook of research on protecting and managing global indigenous knowledge systems*. Hershey, PA: IGI Global.

Olusanya, B.O., Hadders-Algra, M., Breinbauer, C., Williams, A.N., Newton, C.R.J., & Davis, A.C. (2021). The conundrum of a global tool for early childhood development to monitor SDG indicator 4.2.1. *The Lancet Global Health, 9*(5), e586–e587. doi:10.1016/s2214-109x(21)00030-9

Subosa, M., & West, M. (2018). *Re-orienting education management information systems (EMIS) towards inclusive and equitable education and lifelong learning*. Paris: UNESCO.

UNDESA. (2022). *The Sustainable Development Goals report*. New York: United Nations Department of Economic and Social Affairs (UNDESA).

UNESCO. (2019). *Global Proficiency Framework for Reading and Mathematics*. Paris: UNESCO.

UNESCO. (2020). *Global education monitoring report, 2020: Inclusion and education: all means all*. Paris: UNESCO

UNICEF. (2021). *The case for better governance of children's data: A manifesto*. New York: UNICEF.

Unterhalter, E. (2019). The many meanings of quality education: Politics of targets and indicators in SDG4. *Global Policy, 10*(S1), 39–51. doi:https://doi.org/10.1111/1758-5899.12591

World Bank. (2021). *Data for better lives*. Washington, DC: World Bank.

Index