SPRINGER NATURE
Sustainable Development Goals Series
Connecting the Goals

Governing the Sustainable Development Goals

Quantification in Global Public Policy

Justyna Bandola-Gill
Sotiria Grek
Marlee Tichenor

OPEN ACCESS

palgrave
macmillan

Sustainable Development Goals Series

The **Sustainable Development Goals Series** is Springer Nature's inaugural cross-imprint book series that addresses and supports the United Nations' seventeen Sustainable Development Goals. The series fosters comprehensive research focused on these global targets and endeavours to address some of society's greatest grand challenges. The SDGs are inherently multidisciplinary, and they bring people working across different fields together and working towards a common goal. In this spirit, the Sustainable Development Goals series is the first at Springer Nature to publish books under both the Springer and Palgrave Macmillan imprints, bringing the strengths of our imprints together.

The Sustainable Development Goals Series is organized into eighteen subseries: one subseries based around each of the seventeen respective Sustainable Development Goals, and an eighteenth subseries, "Connecting the Goals," which serves as a home for volumes addressing multiple goals or studying the SDGs as a whole. Each subseries is guided by an expert Subseries Advisor with years or decades of experience studying and addressing core components of their respective Goal.

The SDG Series has a remit as broad as the SDGs themselves, and contributions are welcome from scientists, academics, policymakers, and researchers working in fields related to any of the seventeen goals. If you are interested in contributing a monograph or curated volume to the series, please contact the Publishers: Zachary Romano [Springer; zachary.romano@springer.com] and Rachael Ballard [Palgrave Macmillan; rachael.ballard@palgrave.com].

More information about this series at
https://link.springer.com/bookseries/15486

Justyna Bandola-Gill • Sotiria Grek
Marlee Tichenor

Governing the Sustainable Development Goals

Quantification in Global Public Policy

palgrave
macmillan

Justyna Bandola-Gill
University of Edinburgh
Edinburgh, UK

Marlee Tichenor
Durham University
Durham, UK

Sotiria Grek
School of Social and Political Science
University of Edinburgh
Edinburgh, UK

ISSN 2523-3084 ISSN 2523-3092 (electronic)
Sustainable Development Goals Series
ISBN 978-3-031-03937-9 ISBN 978-3-031-03938-6 (eBook)
https://doi.org/10.1007/978-3-031-03938-6

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

This publication is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 715125 METRO (ERC-2016-StG) ('International Organisations and the Rise of a Global Metrological Field', 2017–2022, PI: Sotiria Grek)

Acknowledgements

Every book is a journey; as it unfolds, it is enriched by all the encounters and experiences along the way. This book would never have materialised without the financial, scholarly and administrative support of the European Research Council (ERC-2016-STG METRO project, grant agreement No. 715125); we are immensely grateful for having been supported so generously for the last five years. The ERC support was fundamental not only for conducting the research itself but also for offering us the stability of a five-year period in order to develop the collegiality and shared intellectual agenda that brought this book to fruition.

Second, we would like to extend our gratitude to all the interviewees and organisations that participated in the research and shared their knowledge, views and experiences. Without them, their insights and reflections, we would never have been able to collect the data and do the analysis that has informed the arguments of this book.

We would also like to thank the School of Social and Political Science, University of Edinburgh, and particularly colleagues at the Research Office, who have been supporting this research from the very first day it began: particular thanks go to our colleague Eirini Souri, who gave us valuable project management advice and helped us with the more mundane but no less important aspects of conducting such a large and complex project.

Further, many thanks to the department of Social Policy and colleagues at the Centre for Science, Knowledge and Policy (SKAPE), who commented and advised on the development of this work from inception to delivery. Thanks go to the Edinburgh colleagues Richard Freeman, Marc Geddes, Martin Lawn, Catherine Montgomery, Jarmo de Vries, Agomoni Ganguli-Mitra, Steve Sturdy, Devi Sridhar, Jo Shaw, James Smith and Lindsay Paterson. Many thanks for the continuous support,

collaborations and engagement with our research and ideas to colleagues further afield: Katherine Smith (Strathclyde), Jenny Ozga (Oxford), Paolo Landri (IRPPS, Naples), Toni Verger (Barcelona), Mary Morgan (LSE), Wendy Espeland (Northwestern), Richard Rottenburg (Witwatersrand), Sakiko Fukuda-Parr (New School, New York), Sally Engle-Merry (New York University), Isabel Rocha de Siqueira (PUC-Rio), Johannes M. Waldmüller (FLACSO-Ecuador), Krushil Watene (Massey University, Auckland), Mandy Yap (Australian National University, Sydney), Kate Williams (Melbourne), John Berten (Tübingen), Clara Fontdevila (Glasgow), Anna Pultar (Heriot-Watt), Marissa Mika (University of California, Berkeley), Branwyn Poleykett (University of Amsterdam), Cori Hayden (University of California, Berkeley), Vincanne Adams (University of California, San Francisco). We would also like to thank ex-members of the research team: Matteo Ronzani (University of Manchester), Jennifer Ferguson (University of Stirling) and the METRO visiting fellow Te-Anne Robles (Sheffield). Finally, a very special word of thanks to our colleague Christina Boswell: her friendship, collegiality and support have been a continuous source of strength and inspiration over the years.

Justyna would like to dedicate this book to her family—her husband Nav and their son Rohan—two men who fill every moment with love, joy and adventure. She would also like to thank people who first taught her that learning never stops—Urszula, Bogdan, Joanna and Bartek.

Sotiria would like to thank her partner, Torcail, and their three children, Asterios, Evie and Electra; life without them would not have been half as fun. She would like to dedicate the book to her mum, Roula, and her late father, Yannis: their love, kindness and wisdom have always been a shining light.

Marlee sends thanks and love to her partner—David; her family—LeAnn, Terry, Lauren and Chase; and her chosen family—particularly Zoe and Beth—for bringing meaning, laughter and love to her life.

Praise for *Governing the Sustainable Development Goals*

"This fascinating book addresses how the fundamental logic of quantification drives global public policy. No longer merely a tool of governance, quantification is now deeply embedded as the epistemology of infrastructure. This book is essential reading for those who want to think deeply about how development goals are created and enacted on a global scale."

—Wendy Espeland, *Professor of Sociology, Northwestern University, USA*

"In this well researched investigation of the SDGs, Justyna Bandola-Gill, Sotiria Grek and Marlee Tichenor unpack the complex interplay of actors, measurements, inter-relationships and policy making through indicators. This is an important contribution that advances the study of indicators as they increasingly shape global governance institutions."

—Sakiko Fukuda-Parr, *Professor of International Affairs, The New School, USA*

Contents

About the Authors

Justyna Bandola-Gill is a Research Fellow at the University of Edinburgh (where she obtained her PhD) and Associate Director of SKAPE—Centre for Science, Knowledge and Policy. She is a sociologist of scientific knowledge, working at the intersection of STS and public policy. Her research explores knowledge and expertise as well as quantification and evaluation across different settings, including higher education and sustainable development.

Sotiria Grek is Professor of European and Global Education Governance at the School of Social and Political Science, University of Edinburgh. Sotiria's work focuses on the field of quantification in global public policy, with a specialisation in the policy arenas of education and sustainable development. She is the Principal Investigator of the European Research Council funded project 'International Organisations and the Rise of a Global Metrological Field' (METRO, 2017–2022). She has co-authored (with Martin Lawn) *Europeanising Education: Governing A New Policy Space* (2012) and co-edited (with Joakim Lindgren) *Governing by Inspection* (2015), as well as the *World Yearbook in Education: Accountability and Datafication in Education* (with Christian Maroy and Antoni Verger; 2021).

Marlee Tichenor is a medical anthropologist and a lecturer in the Department of Anthropology at Durham University. She teaches courses on global health, data practices as human practices and postcolonial technoscience. She received her PhD from the University of California, Berkeley, and University of California, San Francisco, and she has written about the fight against malaria in Senegal, the impact of the World Bank on global health and the global goal of universal health coverage.

Acronyms

BMGF	The Bill & Melinda Gates Foundation
CONEVAL	National Council for the Evaluation of Social Development Policy
DESA-Pop Div	The Population Division of the Department of Economic and Social Affairs of the United Nations
DFID	UK's former Department for International Development
EFA	Education for All
EU	European Union
EUROSTAT	European Statistics
ETAG	Extended Technical Advisory Group
HLPF	High-Level Political Forum on Sustainable Development
FGM	Female Genital Mutilation
Global Fund	The Global Fund to Fight AIDS, Tuberculosis and Malaria
GDP	Gross Domestic Product
GAML	Global Alliance for Monitoring Learning
GEMR	Global Education Monitoring Report
GCL	Global Compact for Learning
GPE	Global Partnership in Education
IAEG-MDGs	Inter-Agency and Expert Group on the Millennium Development Goal Indicators
IAEG-SDGs	Inter-Agency and Expert Group on the Sustainable Development Goal Indicators
IEG	The World Bank's Independent Evaluation Group
IFAD	International Fund for Agricultural Development
IFC	International Finance Corporation

ILO	International Labour Organisation
IPA	Interpretive Policy Analysis
IOs	International Organisations
IHME	Institute for Health Metrics and Evaluation
IOM	International Organization for Migration
OECD	Organisation for Economic Co-operation and Development
OPHI	Oxford Poverty and Human Development Initiative
OWG	Open Working Group on Sustainable Development Goals
LMTF	Learning Metrics Taskforce
MDGs	Millennium Development Goals
MPI	Multidimensional Poverty Index
NGOs	Non-governmental organisations
NSO	National Statistical Office
NSS	National Statistical System
NPF	Narrative Policy Framework
PEPFAR	United States President's Emergency Plan for AIDS Relief
PISA	Programme for International Student Assessment
SDG	Sustainable Development Goal
SDMX	Statistical Data and Metadata eXchange
STS	Science and Technology Studies
TAG	Technical Advisory Group
TCG	Technical Cooperation Group
UN	United Nations
UNAIDS	Joint United Nations Programme on HIV/AIDS
UIS	UNESCO Institute of Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
UNDESA	United Nations Department of Economics and Social Affairs
UNDP	United Nations Development Programme
UNFPA	United Nations Population Fund
UNHCR	United Nations Higher Commissioner for Refugees
UNGA	United Nations General Assembly
UNODC	United Nations Office on Drugs and Crime
UNSD	United Nations Statistical Division
UNSC	United Nations Statistical Commission
USAID	United States Agency for International Development
WB	World Bank
WHO	World Health Organisation
VNR	Voluntary National Review

List of Charts

Knowledge Production for the SDGs: Developing the Global Indicators

Scripting the SDGs: The Role of Narratives in Governing by Goals

SDGs and the Rise of an Epistemic Infrastructure: Actors' Networks, Partnerships and Conflicts in the Education SDG

The Sustainable Development Goals as Epistemic Infrastructures

1 More Than Numbers?

In June 2021, the artistic director of the London Biennale, Es Devlin, transformed the courtyard of the Somerset House into the 'Forest for Change': visitors were taken through a journey into the forest, where they walked around to discover the United Nations' Global Goals for Sustainable Development—more commonly known as the SDGs. The SDGs represent the UN's ambitious goal-setting agenda to eradicate poverty, inequality and climate change. According to the Biennale's website, when preparing for the show, Es Devlin was told that trees had been forbidden from the courtyard at Somerset House when the building was originally conceived 250 years ago. Es decided to 'counter this attitude of human dominance over nature, by allowing a forest to overtake the entire courtyard' (London Design Biennale, 2022). In subverting the rules of Somerset House's Enlightenment-era designers, Devlin and her team considered the transformational power of nature to create real change: 'The UN Global Goals offer us clear ways to engage and alter our behaviour and it is our hope that an interaction with the Goals in the forest will be transformative' (London Design Biennale, 2022).

As the planet's climate continues to deteriorate and the world slowly emerges from a lethal global pandemic, the SDGs are indeed representing a transformational moment for humankind and the ways we choose to live in this world. This book focuses on this transformational potential and attempts to understand how and why the SDGs' monitoring agenda is as unique as it has become ubiquitous.

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_1

How have a set of goals and a list of indicators captured the imagination of artists and campaigners, of activists and policymakers? What is it about the SDGs that have come to represent a different way of doing things both in measurement and in policy terms?

Indeed, the SDGs have become a constant feature of our daily lives as they are implemented by a variety of organisations—from governments to universities, the private sector and civil rights organisations. As we will explore in this book, this omnipresence of the SDGs is emblematic not only of the power of this particular initiative but also of the changing nature of the politics of numbers. In analysing the hegemony of the SDGs and the changing nature of numbers that govern, we are confronted with the question: what makes the SDGs any different from existing technocratic measurement tools and thus worthy of detailed analysis and attention? This book offers a view on quantification in public policy that goes beyond looking at its specific tools and effects. We argue that the power of numbers has gone further than that: quantification has become not only a way of *steering* action in global public policy, but rather it has emerged as the key process of *creating* spaces for governing, participation and measurement. As such, quantification has evolved from being the mere 'bricks and mortar' of governance to become its infrastructure itself—or, what this book calls, the 'epistemic infrastructure' of global public policy.

The notion of the epistemic infrastructure aims to capture the political work of numbers in creating connections between actors, constructing new frameworks of thinking and doing in policy and ultimately becoming the carrier of a new governing paradigm. The infrastructural lens on measurement in global public policy allows us to explore not only the metrics themselves but also the socio-political environments which enable their political effects. As such, the focus here moves from the well-established idea of 'governing by numbers' (Miller, 2001; Scott, 1998) to the focus on 'governing numbers' and the 'governing *of* numbers' as a key mode of producing policy knowledge and a unified global space to govern. Numbers do not merely influence the knowledge that governs action; rather, as we will show, quantification emerges as a new global public policy paradigm that shapes and reshapes the very architecture of transnational governance. As we will show, in recent decades, this new architecture of transnational governance has coalesced around the concepts of 'sustainability' and 'sustainable development'.

2 Sustainable Development and the Rise of Sustainability Politics

How did it come about that 'sustainability' and 'sustainable development' became the central governing principles of the 'first truly global policy agenda', the SDGs (UNGA, 2015)? As this book will show, in both explicit and implicit ways, the SDGs built on and departed significantly in substance and form from the Millennium Development Goals (MDGs). The MDGs—agreed upon by 189 UN member states with the Millennium Declaration in 2000—and their 60 indicators were widely critiqued as driven by a small number of powerful political entities (the United States, Europe and Japan) in order to effect change exclusively in poor countries (Amin, 2006; Saith, 2006). For 15 years, the MDGs largely defined development priorities for multilateral, bilateral and philanthropic organisations in the Global South, creating a donor-led vision of global progress. For this reason, although the MDGs were widely heralded for highlighting poverty reduction and social development as the most important development problems of the new millennium, they were also criticised for being 'reductionist' and for framing 'development as a top-down approach to meeting basic needs, promoting a target driven strategy, and [de-contextualizing] from local settings' (Fukuda-Parr & McNeill, 2019, p. 8).

On the contrary, from the beginning, the 2030 Agenda with its Sustainable Development Goals was designed to be country driven. With this priority in mind, the SDGs were developed through two parallel consultancy processes, the Secretary General-run 'Post-2015 Development Agenda' and the Open Working Group (OWG), that emerged from the Rio+20 Conference on the Environment and Development. The 'MDG plus' that was to emerge from the UN-led 'Post-2015 Development Agenda' was conceived to be much the same as that which came before, while the OWG was much more revolutionary in the 'structural change' of the country-led coalition, which called for working towards issues of 'poverty, environmental sustainability, economic development, and social equity'—going far beyond a focus on 'basic needs' (Fukuda-Parr & McNeill, 2019, p. 9). Thus, it became very important that the SDGs were country-led, and that the UN agencies that had been in the driver's seat for the MDGs give up the wheel for countries to lead the process. Fundamentally, the SDGs replaced the MDGs' poverty agenda with an agenda for 'sustainable development' (Fukuda-Parr, 2016).

Productively, in the past few decades in the UN space, 'sustainability' and 'sustainable development' have become generatively vague terms. There is no question that the SDGs have their origin in the Rio+20 conference in 2012, meant to mark the 20th anniversary of the 1992 UN Conference on Environment and

Development—the 'Earth Summit' (Dodds et al., 2012; Dodds et al., 2017). The Rio Earth Summit was the culmination of a movement towards global attention towards and prioritisation of protecting the environment from *unsustainable* modes of economic development, first codified in 1987 in a report by the United Nations World Commission on Environment and Development (WCED), also known as the Brundtland Report or 'Our Common Future'. In her forward to the report, Chair Gro Brundtland[1] emphasised that the environment had to be accompanied with a rethinking of the concept of development. Development could not be conceived of as 'what poor nations should do to become richer' when it was clear that 'many of the development paths of the industrialized nations are clearly unsustainable' (WCED, 1987, p. 7), thereby introducing the concept of 'sustainable development' to solve this quandary.

Some authors argue that the goal of sustainable development, as defined in this highly influential text, was always to 'reconcile the conflicting goals of environmental protection and economic growth' and proclaim a 'utopia of a society where no obvious concessions are necessary' (Quental et al., 2011, p. 16). Conversely, Gasper et al. (2019) argue that the concept of sustainable development has changed in subtle but important ways since the Brundtland Report, which followed in the path set out by the first UN Conference on the Human Environment in 1972. While the authors of the Brundtland Report argued that 'growth was necessary to reduce global poverty', they also emphasised 'the imbalance between consumption patterns of the wealthy and the poor', making space for consumption to be a key issue to addressing sustainable development (Gasper et al., 2019, p. 84). However, over the decades since the report, the issue of the volume of consumption became less of a standalone issue and was increasingly conceptualised alongside production, with the ultimate conclusion that protecting 'the environment and aspiring to unendingly higher levels of consumption for everyone are not seen as contradictory' (2019, p. 85).

Irrespective of whether this ambiguity about what counts as 'sustainable' existed from the origins of 'sustainable development' in the Brundtland Report or not, the concept of sustainable development is certainly a highly malleable one.

[1] Brundtland had been the Prime Minister of Norway, and the Minister of Environment before that, although she was a public health scholar by training. She credits her academic training in public health and political training in the Ministry of the Environment for the alchemy that led to the concept of sustainable development as it appears in the Brundtland Report: 'by that reasoning, we had more breakthrough in public debate in Norway. Because I linked the environment to people, and to health, and to the future of our societies' (Brundtland, 2011). This extended the points made at the first UN conference on the environment, the 1972 UN Conference on the Human Environment.

Fred Gale argues that 'not only does sustainability encompass economic, social and environmental components but that, in seeing to integrate these components into existing approaches, political economists of all persuasions interpreted it as compatible with their established conceptions' (Gale, 2018, p. x). Ironically, for example, 'sustainability' can come to be compatible with 'sustained economic growth' for neoclassical economists (Aznar-Márquez & Ruiz-Tamarit, 2016), while also be flexible enough to be used in environmental activist discourse that fights against such conceptions of growth. As we will see throughout the book, this ambiguity about what counts as 'sustainable' placed the responsibility of actualising ambitious goals about thinking concurrently about the environment, the social and the economic—the three pillars of the SDG agenda—on the shoulders of national and UN statisticians, who had to concretise the SDGs' goals and targets into measurable indicators. But how come measurement has acquired such a central positioning in the production of global public policy? The next section will grapple with this issue before moving on to discuss the notion of 'epistemic infrastructures' and what they entail.

3 Quantification in Global Public Policy: From Governing by Numbers to Epistemic Infrastructures of Measurement

Sociologists and anthropologists of quantification observed that 'our lives are increasingly governed by—and through—numbers, indicators, algorithms and audits and the ever-present concerns with the management of risk' (Shore & Wright, 2015; 23; see also influential work by Merry, 2011; Sauder & Espeland, 2009; Strathern, 2000). Perhaps unsurprisingly, the literature on the history, politics and social effects of quantification has burgeoned over the last decades (e.g. Alonso & Starr, 1987; Desrosières, 1998; Espeland & Stevens, 2008; Hacking, 1990, 2007; Porter, 1995; Power, 1997; Rose, 1999). Global governance is particularly susceptible to the 'seductions' of quantification (Merry, 2016), as this field relies on the availability of easily comparable and universalising forms of knowledge (Rottenburg & Merry, 2015). Hansen and Porter (2017) suggest that, although it took scholars a long time to recognise the constitutive nature of discourse, we are now well aware of the role of language in shaping reality. However, they suggest that numbers are characterised by additional qualities that make their influence much more pervasive than words. These elements are order, mobility, stability, combinability and precision. Numbers transform complex issues into readily auditable objects (Power, 1997) that are subject to political rationalities (Miller, 2001).

The predominant focus of this vast body of scholarship is on analysing quantification as a set of tools that incites different institutional and political responses (Espeland & Sauder, 2007; Espeland & Stevens, 2008). This focus on the specific effects that metrics ought to produce leads to processes of 'reactivity' (Espeland & Sauder, 2007) or even 'gaming' in order to fit the expected organisational and political scripts for action (Bevan & Hood, 2006; Strathern, 1997). More recently, there is a growing recognition that the power of quantification goes beyond its effectiveness and ubiquity as a set of tools to govern. Quantification is seen instead as a 'logic' (Chun & Sauder, 2021) or a 'culture' (Mennicken & Espeland, 2019; Merry, 2016) on which actors draw in different institutional and bureaucratic settings. As such, quantification is a carrier of broader political, social, cultural and institutional orders. Therefore, quantification is more than just numbers—rather, it is the central machinery of promoting specific modes of governing, such as evidence-based governance, and as such is inherently paradoxical as a logic of governance, as argued by Merry (2016, p. 11):

> Governance by indicators can increase egalitarian decision making and accountability by opening up the basis for decisions to public scrutiny. On the other hand, it can also reinforce inequality and evoke resistance among the governed.

This book follows this line of inquiry by exploring the processes and practices of quantification through the analytical lens of epistemic infrastructure. A social theory interest in infrastructures first emerged in the Science and Technology Studies (STS) literature (Bowker, 1995; Star & Ruhleder, 1996) in order to describe the mix of materials, practices and meanings that comprise interlinked knowledge structures, generating effects and structuring social relations. In the context of sustainable development, the concept of epistemic infrastructures is particularly useful for capturing the emergence, processes and consequences of the ways in which quantification has scaled up, linking different sites of calculation and governance. This analytical lens allows for theorising quantification as a meta-level phenomenon that governs not only through its explicit political effects but rather through creating structures, connections and interdependencies, thus allowing for new governing spaces to emerge.

At the same time, even though the concept of 'epistemic infrastructures' is becoming increasingly prevalent in global public policy literature (e.g., Bueger, 2015), it is often used in a vague and under-theorised way. Building on our use of the term elsewhere (Tichenor et al., 2022), this book aims to unpack this term by proposing a theorisation of an epistemic infrastructure of measurement within the SDGs as an interplay of three levels: the *materialities* of measurement, *interlink-*

ages between actors and measures and finally *new paradigms* of doing global public policy.

The materialities of measurement are the building blocks of the epistemic infrastructures. Just as the physical infrastructures are built from bricks, metal and concrete, epistemic infrastructures are constructed with data, indicators, surveys, reports, data visualisations, etc. Within the SDGs, these materialities entail the complex system of goals, targets and indicators that allow for constructing the concept of 'sustainable development' in practice. This epistemic infrastructure did not emerge all at once; rather, it was built on the existing foundations—the Millennium Development Goals as well as the decades-old statistical systems of the UN countries. Within the SDGs, these materialities were linked together into a framework in the processes of negotiation and governing of the indicator system (as we show in Chap. 2). These fragmented assemblages of different measures and approaches required a process that would unify this wide variety of practices into a coherent global measurement—and consequently governing—programme. Here, the practices of harmonisation of data and metrics take the central stage as the key process of creating sustainable development as a global public policy programme. Through processes of harmonisation (as we discuss in Chap. 3), data and metrics produced by various means and across countries and institutions are transformed into global data and metrics—ones that allow for comparison, benchmarking and—more often than not—competition between countries. Importantly, numbers in the SDGs do not speak for themselves: instead, their meaning is built in context—and this meaning-making process occurs through narrativisation (Chap. 4). Numbers are transformed into political entities—ones that carry values, priorities and ideas that travel across different global and local communities—through the process of storytelling and narrative-making.

The second order of the infrastructure is the **interlinkages** through which these diverse materialities are connected and held together. Here, the central role is played by the epistemic communities, communities of practice and varied networks of experts. The key difference between the MDGs and the SDGs is the different approach to by governing numbers—the MDGs were a predominantly top-down process orchestrated by International Organisations. The SDGs (as we show in Chap. 5) were from the outset designed as a participatory programme with the country members leading the way. Consequently, the process of *governing by numbers* requires a large dose of negotiation and navigation of these emergent communities and networks. The epistemic infrastructure is maintained through interdependencies between these various actors (which we discuss in Chap. 6). Fuelled by both competition and collaboration, International Organisations configure and reconfigure epistemic communities around particular policy arenas, ex-

tending the terrain of the SDGs and thus of global public policy itself in the process. Further, the establishment and maintaining of the infrastructure through the interlinkages requires new forms of expertise. The experts working in this field are no longer merely statisticians and data scientists: increasingly what is needed is a new form of experts—expert brokers (Chap. 7) whose main role is creating and managing the connections between different disparate groups of political actors—the governments, the National Statistical Offices, International Organisations, the third sector and community groups, academia, etc. The proclaimed aim here is for the process of producing numbers to not only represent a technocratic exercise in measurement but also facilitate a democratic process of negotiating the common epistemic order established through quantification. This effort, that is, to inject the technical process with political clout and give it democratic credentials, is at the heart of building the epistemic infrastructure and hence the core analytical question for this book: in other words, what happens when the production of numbers is proclaimed and used as the key venue for democratic decision-making amongst global leaders in their efforts to steer the future?

This book examines this new status quo, facilitated by the SDGs, as the foundation of a **new global public policy paradigm**. In the process of collecting the data for this book, one of our interviewees stated—somewhat provocatively—'there is no global policy, apart from the climate policy'. As we will argue in this book—the SDGs have not only generated a global public policy programme by challenging the ways we think about what global policy is and how it is practised; they have also created the measurement and governing architecture to bring it to fruition. The SDGs are the new paradigm for two reasons: first, they have reshaped the idea of 'sustainable development', making it an 'all encapsulating' concept that unites nearly all policy fields; second, measurement has become the central way of thinking and doing sustainability.

Central to this book is the key role of the production of quantification as one of the primary tools of governing the transnational. The book builds on the literature on the making of measurement infrastructures (Merry, 2019) to argue for the rise of global public policy as an *epistemic infrastructure*: as it will become evident in the chapters of this book, we see global public policy as not merely the outcome of 'governing by numbers' (Miller, 2001). The book moves beyond a theorisation of global policy as the top-down steering of policy action at the national level through the application of soft governing tools, such as indicators and benchmarking. As we will show, instead of examining policy change as the side-effect of measurement processes, we will show how the epistemic infrastructure of the SDGs has become a crucial site for global public policy work: as the measurement space opens up to become an arena of deliberation and negotiation about policy goals and

policy prioritisation, numbers do not merely count; they represent and embody policy directions and become the key venue of policy contestation and consensus.

4 Global Public Policy: A Fluid Concept and a Contested Terrain

The previous section discussed the emergence of global public policy as an epistemic infrastructure. Nonetheless, what does policymaking at the level of the global mean? For Ramesh Thakur and Thomas Weiss (2009), policy refers to the statement of principles and actions that an organisation is likely to pursue in the event of particular contingencies (2009, p. 19). It is different from norms and institutions, in that it is issue-driven: for example, in the question whether 'UN policy' exists, or indeed in whether the SDGs represent a policy framework or not, Evans and Newnham suggest that policy is not simply a set of governing principles, but reflects '…the decision to embark upon certain programmes of action (or inaction) in order to achieve desired goals' (1998, p. 440). Following this definition, 'policymakers' are actors participating in such processes. Yet, a sharp distinction is often made between national/domestic and foreign policy, the latter being the lens most often used for understanding the participation of national actors in intergovernmental decision-making (Thakur & Weiss, 2009).

Diane Stone (2008) contributed substantially to the discussion of the nature and function of global public policy, by questioning the role of states as key actors in policy formulation; she suggested that global public policy is a multi-centric, transformative, complex global political system with multiple issue-regimes to govern contemporary global challenges. She emphasised the need to re-conceptualise the global public policy space as a global agora, to pay greater attention to the interactions between public and private actors as well as the role of knowledge producers in shaping the field. In a similar vein, Reinicke (1998) saw the role of networks as key in the production of global public policy in that they are effective at bringing together different groups of actors and finding common solutions to common problems. According to him, global public policy networks 'govern without governments'. These networks achieve this 'by placing new issues on the global agenda; negotiating and setting global standards; gathering and disseminating knowledge; making new markets where lacking or deepening markets that are failing; and innovating implementation mechanisms for traditional intergovernmental treaties' (Reinicke, 1998, p. xv).

Indeed, most scholars of global public policy agree with the premise that the world is increasingly globalised and interdependent (Nagel, 1990; Soroos, 1986;

Stone, 2008). Benner et al. (2003) write that 'since the early 1990s, the driving forces of globalisation, technological change and economic and political liberalization have fundamentally transformed conditions for effective and legitimate governance' (2003, p. 18). More importantly, scholars argue for the significance of studying global public policy not as an add-on to national policymaking, but as an important space where governing decisions are taken, decisions that affect national politics, too:

> An agenda of global problems can be identified. Elements of an international policy process have been in place for at least several decades. Policies containing regulations and programs have been incorporated into treaties and resolutions. Finally, steps have been taken to implement and review the policies that have been adopted. The nature of contemporary world politics cannot be adequately understood without knowledge of these cooperative efforts at global problem solving. (Soroos, 1986, p. 374)

To date, literature has been dominated by a range of approaches for understanding the production of global public policy. According to Brinkerhoff, the *partnership approach* refers to 'a dynamic relationship among diverse actors, based on mutually agreed objectives, pursued through a shared understanding of the most rational division of labour based on the respective comparative advantages of each partner' (Brinkerhoff, 2002, p. 325). On the other hand, perhaps the most common and dominant way of studying global public policy has been through a focus on *network formation*:

> Global public policy networks build bridges across different sectors and levels, bringing together actors from governments, international organizations, civil society, and business…Unlike traditional hierarchical organizations, these networks are evolutionary in character and flexible in structure. They bring together disparate groups with oftentimes considerably varying perspectives, combining knowledge from different sources in new ways to result in new knowledge. (Benner et al., 2003, p. 18)

Related to the rich literature to the rise of networks in global governance is the study of transnational advocacy networks which coordinate action around a 'principled issue'; they create links and rally around the convergence of a range of actors, from civil society, to states, IOs, global philanthropists, business and others in creating momentum for policy agenda setting on key global issues (Keck & Sikkink, 1998). International relations theorists have written persuasively about the rise of *regimes* in global public policy, and the ways that different regimes, such as human rights, humanitarianism, development and security, overlap and intersect in order to create a governing complex that requires inter-institutional cooperation at the transnational level (Betts, 2009). Finally, a key approach to the

understanding of global public policy is through the emergence of *norms*: Thakur and Weiss suggest that norm-setting is a key function of global public policy, since 'if policy is to escape the trap of being ad hoc, episodic, judgemental, and idiosyncratic, it must be housed within an institutional context' (2009, p. 20). As we outline in the book network formation and practices of expert brokering are crucial for constituting the global public policy of the SDGs and constitute the second-order level of the epistemic infrastructure. We build on this existing rich literature of global public policy by focusing on how the governing of numbers in these global spaces has become the privileged mode of producing unified, concretised policy on the global level.

5 Research Design

This analysis has been based on a project funded by the European Research Council (grant number: 715125, Principal Investigator: Sotiria Grek), 'International Organisations and the Rise of a Global Metrological Field' (or METRO for short). The project draws on a rich set of mixed-methods data, including analysis of documents, semi-structured interviews with the key experts in International Organisations and Social Network Analysis of meetings. Overall, this book draws on a rich dataset of over 80 interviews with key experts in these epistemic communities, as well as the careful analysis of documents, including flagship reports, policy and strategic documents (such as declarations, position papers and action plans), internal documents produced by IOs (including meeting agendas, open consultations and PowerPoint presentations) and research articles published by actors in these networks.

The research design was grounded in a comparative case study of different policy fields, examining the SDGs as a whole, but also focusing deeper on the cases of education (in particular SDG 4), poverty (SDG 1) and statistical capacity development (cutting across all the SDGs). In the examination of the SDGs as a whole, we also conducted document analysis and interviews around the policy fields of health, migration and sustainable tourism.

First, in relation to the education goal (SDG4), it was produced in a context of increased datafication in education governance as the prime mode of knowing and reforming complex education systems around the world. The rise of large international assessments created a wealth of statistical information and thus allowed states and transnational agencies for the first time to construct comparative knowledge about education performance. Thus, SDG4 promises to 'ensure inclusive and equitable quality education and promote lifelong learning opportunities for all'

(United Nations General Assembly, 2015, p. 14). The SDG4 represents the single biggest attempt to bring together a vast array of actors and countries in order to construct universal education indicators, as well as to decide on the appropriate methodologies and data sources. Like all SDGs, it is a country-led, global exercise—led by UNESCO but with the collaboration and close involvement of all major International Organisations (IOs).

Through an in-depth analysis of texts and interviews, this case study explores the conundrum of securing accountability of this global performance monitoring project through ensuring the objective validity of its measurement tools, whilst promoting the democratic and equal participation of all actors. UNESCO, as the custodian agency of SDG4, has a double accountability obligation to participating countries: firstly, the robust and objective monitoring of progress towards the SDG4 goals, and secondly, the participatory and democratic, equitable process in which all member countries have a voice and stake in the project. As a result, although the UNESCO Institute of Statistics has been significantly reinvigorated in relation to its statistical capacity, it has also put great emphasis on the participatory, inclusive and consensual aspects of the agenda.

Secondly, the project focused on the ending poverty goal (SDG1) as one of the key challenges of sustainable development. The realisation of this goal is most commonly—both discursively and materially—linked to the production of high-quality poverty knowledge. At the same time, the quantification of poverty knowledge is strongly contested. The UNICEF Innocenti report describes the measurement of poverty as a 'crisis in monitoring' (2015). Indeed, there has been profound disagreement and controversy around the measurement systems of poverty—both in the academic and policy worlds. One of the factors accelerating this crisis is the increase in the number of approaches to measurement promoted by International Organisations. Just in the last twenty years, the number of global measures of poverty increased from one (the popular dollar-per-day International Poverty Line introduced by the World Bank) to eight different monetary and multidimensional approaches.

This case explores—through document analysis and semi-structured interviews—the dynamics of poverty governance in the situation in which multiple measurement approaches compete. In the face of the multiplicity of different measures, International Organisations employ various strategies to assess, create and communicate the epistemic, political and strategic values of poverty indicators. Consequently, the process of measurement—and the controversies around it—is a domain of navigating different legitimating forces.

Last but not least, METRO examined statistical capacity building in depth, as representing one of the SDGs (SDG17), but also being the key driver of change in regard to the development of the agenda as a whole. In the process of monitoring the MDGs, the lack of data in many countries or sub-national regions was highlighted as a problem that development agencies must put on their agendas. This lack of official data was particularly stark in the face of the rapidly changing technology landscape that has led to a 'data revolution' in many parts of the world, which has constructed elaborate alternate data and meta-data collection systems alongside official statistical systems. With these needs and inequalities in mind, United Nations member states and IOs put emphasis on the development of statistical capacity in the Global South and incorporated it as an indicator for monitoring within SDG17. Simultaneously, statistical capacity development is also presented as necessary for the functioning of the global sustainable development agenda in its entirety, maintaining and creating the infrastructure for the 231 unique indicators to be monitored by custodian UN agencies.

Based on interviews with key figures within IOs, network analysis of advocates of statistical capacity development, and critical discourse analysis of key texts, we investigated the debates and processes of developing global consensus on principles and standards for statistics, statistical systems, and their development, including in the part of the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs), the UN Statistics Division, and others. At the heart of these debates is the tension between the 'empowering' or 'democratic' nature of data—to 'make people count'—and the work of creating universal standards for measurement, as well as tensions between different practices of statistical estimation and representation.

6 Book Summary

The book begins with an attention to the production of indicators and their harmonisation for the SDG framework in Chaps. 2 and 3: governed by carefully defined protocols and networks, the SDGs' 231 unique indications have been deliberated, chosen and refined for inclusion, making each indicator a microcosm of the knowledge and policy practices that fuel the epistemic infrastructure as a whole. Chapter 4 moves on to discuss the role of narrative-making in the making of the epistemic infrastructure: through the intertwinement of numerical, discursive and visual narratives, we show the significance of story-making for giving the numbers 'heart and soul'.

The book proceeds to focus on the interlinkages of actors, materialities and processes in the production of the epistemic infrastructure: Chap. 5 focuses on the tense politics of ensuring democratic and technocratic accountability, and thus shows the work of numbers in producing venues for policy production. Chapter 6 moves on to examine the role of networks and meetings: it shows the generative potential of conflict and failing metrics as a way of keeping the infrastructure going, always expanding and moving into new territories. Chapter 7 examines the politics of producing expertise in such a complex, sometimes even chaotic, field: the central role of the experts in IOs in the process of governance of the SDGs lies not solely in providing technical guidance but rather in the mediation and brokerage between actors and fields. As such, the legitimacy and effectiveness of experts rely on their ability to mediate connections, create and communicate common meanings of problems and integrate multiple bodies of knowledge.

Finally, we conclude the book (Chap. 8) with a discussion of the rise of global public policy as an epistemic infrastructure: we show the ways that the SDGs as a monitoring and governing agenda have transformed the role of quantification, not as merely the facilitator and enabler of policy decisions taken elsewhere, but as the prime site of governing the future itself.

References

Alonso, W., & Starr, P. (1987). *The politics of numbers*. Russell Sage Foundation.

Amin, S. (2006). The millennium development goals: A critique from the South. *Monthly Review, 57*(10), 1.

Aznar-Márquez, J., & Ruiz-Tamarit, J. R. (2016). Environmental pollution, sustained growth, and sufficient conditions for sustainable development. *Economic Modelling, 54*, 439–449.

Benner, T., Reinicke, W., & Witte, J. M. (2003). Global public policy networks: Lessons learned and challenges ahead. *The Brookings Review, 21*(2), 18–21.

Betts, A. (2009). *Forced migration and global politics*. Wiley-Blackwell.

Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration, 84*(3), 517–538.

Bowker, G. C. (1995). Second nature once removed: Time, space and representations. *Time & Society, 4*(1), 47–66.

Brinkerhoff, J. M. (2002). Global public policy, partnership, and the case of the world commission on dams. *Public Administration Review, 62*(3), 324–336.

Brundtland, G. (2011). Gro Harlem Brundtland, Former World Health Organization Director-General. Voices in Leadership. Retrieved September 16, 2021, https://www.hsph.harvard.edu/voices/events/brundtland/

Bueger, C. (2015). Making things known: Epistemic practices, the United Nations, and the translation of piracy. *International Political Sociology, 9*(1), 1–18.

Chun, H., & Sauder, M. (2021). The logic of quantification: Institutionalizing numerical thinking. *Theory and Society, 51*, 335–370.
Desrosières, A. (1998). *The politics of large numbers: A history of statistical reasoning*. Harvard University Press.
Dodds, F., Donoghue, A. D., & Roesch, J. L. (2017). *Negotiating the sustainable development goals*. Routledge.
Dodds, F., Strauss, M., & with Strong, M. F.. (2012). *Only one Earth: The long road via Rio to sustainable development*. London: Routledge.
Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology, 113*(1), 1–40.
Espeland, W. N., & Stevens, M. L. (2008). A sociology of quantification. *European Journal of Sociology, 49*(3), 401–436.
Evans. G., & Newnham, J. (1998). *Dictionary of International Relations*. London: Penguin.
Fukuda-Parr, S. (2016). From the millennium development goals to the sustainable development goals: Shifts in purpose, concept, and politics of global goal setting for development. *Gender & Development, 24*(1), 43–52.
Fukuda-Parr, S., & McNeill, D. (2019). Knowledge and politics in setting and measuring the SDGs: Introduction to special issue. *Global Policy, 10*(S1), 5–15.
Gale, F. P. (2018). *The political economy of sustainability*. Edward Elgar Publishing.
Gasper, D., Shah, A., & Tankha, S. (2019). The framing of sustainable consumption and production in SDG 12. *Global Policy, 10*(S1), 83–95.
Hacking, I. (1990). *The taming of chance*. Cambridge University Press.
Hacking, I. (2007). *Proceedings of the British Academy, Volume 151, 2006 lectures*. British Academy.
Hansen, H., & Porter, T. (2017). What do big data do in global governance? *Global Governance, 23*(1), 31–42.
Keck, M., & Sikkink, K. (1998). *Activists beyond borders: Advocacy networks in international politics*. Cornell University Press.
London Design Biennale. (2022). 'The Global Goals Pavilion: Forest for Change'. At https://www.londondesignbiennale.com/participant/the-global-goals-pavilion [19.5.2022].
Mennicken, A., & Espeland, W. N. (2019). What's new with numbers? Sociological approaches to the study of quantification. *Annual Review of Sociology, 45*(1), 223–245.
Merry, S. E. (2011). Measuring the World: Indicators, human rights, and global governance. *Current Anthropology, 52*(S3), S83–S95.
Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. University of Chicago Press.
Merry, S. E. (2019). The sustainable development goals confront the infrastructure of measurement. *Global Policy, 10*(S1), 146–148.
Miller, P. (2001). Governing by numbers: Why calculative practices matter. *Social Research, 68*(2), 379–396.
Nagel, S. (1990). Introduction to global policy studies. *International Political Science Review, 11*(3), 303–306.

Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Power, M. (1997). *The audit society: Rituals of verification*. OUP.

Quental, N., Lourenço, J. M., Nunes, F., & da Silva. (2011). Sustainable development policy: Goals, targets and political cycles. *Sustainable Development, 19*(1), 15–29.

Reinicke, W. (1998). *Global public policy: Governing without Government?* Brookings Institution Press.

Rose, N. (1999). *Powers of freedom: Reframing political thought*. Cambridge University Press.

Rottenburg, R., & Merry, S. E. (2015). A World of indicators: The making of governmental knowledge through quantification. In R. Rottenburg, S. E. Merry, S.-J. Park, & J. Mugler (Eds.), *The World of Indicators* (pp. 1–33). Cambridge University Press.

Saith, A. (2006). From universal values to millennium development goals: Lost in translation. *Development and Change, 37*(6), 1167–1199.

Sauder, M., & Espeland, W. N. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review, 74*(1), 63–82. https://doi.org/10.1177/000312240907400104

Scott. (1998). *Seeing like a State*. Yale: Yale University Press.

Shore, C., & Wright, S. (2015). Governing by numbers: Audit culture, rankings and the new World order. *Social Anthropology, 23*(1), 22–28. https://doi.org/10.1111/1469-8676.12098

Soroos, M. (1986). *Beyond sovereignty: The challenge of global sovereignty*. University of South Carolina Press.

Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research, 7*(1), 25.

Stone, D. (2008). Global public policy, transnational policy communities, and their networks. *Policy Studies Journal, 36*(1).

Strathern, M. (1997). "Improving Ratings": Audit in the British university system. *European Review, 5*(3), 305–321.

Strathern, M. (Ed.). (2000). *Audit cultures: Anthropological studies in accountability, ethics and the academy*. Routledge.

Thakur, R., & Weiss, T. (2009). United nations 'Policy': An argument with three illustrations. *International Studies Perspectives, 10*(1).

Tichenor, M., Engle, S. M., Grek, S., & Bandola-Gill, J. (2022). Global public policy in a quantified World: Sustainable development goals as epistemic infrastructures. *Policy & Society*. https://doi.org/10.1093/polsoc/puac015

United Nations Children's Fund (UNICEF). (2015). Measuring child poverty: New league tables of child poverty in the world's rich countries. Retrieved October 20, 2015, from http://www.unicef-irc.org/publications/pdf/rc10_eng.pdf

United Nations General Assembly (UNGA). (2015). *Transforming our world: The 2030 Agenda for Sustainable Development (A/ RES/70/1)*. United Nations.

WCED. (1987). *Report of the world commission on environment and development: Our common future*. World Commission on Environment and Development.

Knowledge Production for the SDGs: Developing the Global Indicators

1 Introduction

The most important entry point for understanding how the SDGs produce global public policy is an examination of their indicators: the ways they are deliberated, chosen, refined and measured on the global level, which is the key venue for producing quantified global governing knowledge. In contrast to the way that the MDGs' 60 indicators were chosen, the United Nations Statistical Commission was tasked with creating a system for choosing and refining indicators for inclusion in the SDGs that would be led by member states rather than International Organisations. Such an innovation in the governing of the monitoring agenda was seen as allowing for greater refinement, since both methodologies and data sources were about to expand and change over the period 2015–2030. The UNSC's designated working group for this deliberative work, the Inter-Agency and Expert Group for the Sustainable Development Goal Indicators (IAEG-SDGs), became the key agency for establishing protocols for evaluating methodologies for producing data for each indicator.

This chapter will analyse the process by which indicators are deliberated, chosen and refined for inclusion in the global SDG monitoring framework. This includes a careful analysis of the evaluative framework for indicators, which is an ongoing process as the SDG framework is seen to be a living document, to allow for further refinement as we ease towards 2030. This evaluative framework is central to the production of the SDGs as a whole: it encapsulates the tensions within

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_2

this 'neutral' space of statistical decision-making and shows us how entwined technical and political accountability are within the global agenda of the SDGs. As such, the evaluation of measurement can be seen as a governing practice. In the next section, we show where this argument sits within the literature on the production of knowledge for governance and the role of knowledge evaluation in governing paradigms. The following section provides a history of the development of the SDGs' tier system, shows how it is a performative—in many senses of the term—space and outlines how it has served both as an engine and an obstacle for certain policy issues, as inclusion within the SDG framework has become one of the key modes of policy advocacy.

2 Knowledge for Governance, Practices of Evaluating Governing Knowledge, and Technical-Political Accountability

Social scientists have shown how the production of quantified knowledge for policy has become a hallmark of contemporary governance. Policy studies scholarship has investigated the nature of evidence-based policymaking and the so-called evidentiary turn in public policy (Duffy, 2017; Miller, 2001; Taylor, 2005). 'The promise of evidence-based policy-making'—as anthropologists Rottenburg and Merry (2015, p. 1) argued—'is that it is not only more objective and less prone to misuse, but also more transparent, more democratic, and more open to public debate than decisions taken by politicians and business leaders with reference to qualitative ways of knowing'. As 'globally circulating knowledge technolog[ies] that can be used to quantify, compare and rank virtually any complex field of human affairs' (Rottenburg & Merry, 2015, p. 5), indicators are key to supporting this promise.

The object of study for this chapter is the evaluation process for indicators' inclusion, classification and refinement in the SDG framework. Our fundamental argument here is that substantive decisions in the space of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs)—which explicitly labels itself as 'neutral' and 'apolitical'—are vital to the global governing paradigm of the SDGs. In this way, the UNSC and the IAEG-SDGs have created a taxonomy of taxonomies through the tier system, as they classify and reclassify the methodologies for measuring economic, social and environmental phenomena on the global level. Informed by Science and Technology Studies (STS) scholarship, we are interested in opening up the practices of evaluation of quantified knowledge that are at the centre of this work. As we will show, this evaluation of indicators includes more than merely the evaluation of methodologies for

measuring phenomena, but indeed entangle considerations about finance, power and alignment of multitudes of policy agendas on national and global levels. For example, in the first meeting of the IAEG-SDGs in June 2015, the representative for the Cameroonian NSO—standing in for 'the African group of countries' represented at the meeting—argued the 'need to establish the costing structure'; second, he emphasised country ownership; and finally suggested the need for a careful elaboration of the process for choosing proxies and making sure indicators aligned with national development plans (IAEG-SDGs, 2015, p. 6).

These evaluatory practices of quantified knowledge production are by definition both technical and political. Building off the 'success' of the MDGs to entangle global policy agendas with quantified knowledge, the SDGs entangle technical and political accountabilities (Bandola-Gill, 2021; Fontdevila & Grek, 2021). As outlined in the 2011 Busan Action Plan, 'reliable and accessible statistics provide the evidence needed to improve decision making, document results, and heighten public accountability' (PARIS21, 2011, p. 2). Political accountability implies a relationship of responsibility, that a governing power—whether a nation-state or a supra-national organisation—is to account for its actions, which impact its citizenry or beneficiary population. As many scholars of quantification have shown, however, in recent decades, accountability—as a form of responsible governance—has become closely tied to quantification. Espeland and Vannebo (2007, p. 22) discuss how, understood 'as creating responsible people and accessible, responsive institutions, accountability is obviously a desirable goal'. However, with the new 'technologies of audit and accountability', discussed in the introduction, came 'new forms of governance and power' (Shore & Wright, 2004, p. 57). In the context of the SDGs, some have expressed anxiety that these practices of counting and evaluating quantitative knowledge, through the introduction and monitoring of indicators on policy issues like gender equality, have themselves effectively become proxies for 'substantive contestation on key policy issues and meaningful accountability mechanisms' (Razavi, 2019, p. 149).

3 Processes and Institutions: Producing Indicators for the SDGs

3.1 Evaluating Statistical Knowledge: Developing Protocols for Choosing Indicators

Central to the production of knowledge for governance in the context of the SDGs is an intricate evaluative model for deliberating, choosing and refining indicators for monitoring progress on the larger framework's targets and goals. This evalua-

tion of quantified knowledge for the SDGs depends on a tripartite system: first, the production of a classification system for legitimizing global indicators for inclusion within the SDG framework (the tier system); second, the relevant protocols for guiding the promotion—or rejection—of different indicators through the said classification system (evidence production and methodology refinement) and third, a network of actors with delineated authority over the deliberation process (the Inter-Agency and Expert Group for the Sustainable Development Goal Indicators, or the IAEG-SDGs). In addition to the above, as we will see below, many indicators follow certain 'path dependencies', carrying with them long-standing epistemic communities who have spent years, or even decades, debating and producing bodies of evidence to support particular ways of measuring phenomena in public policy.

Further, as part of the larger emphasis that the SDGs be much more participatory, national statistics offices (NSOs) and institutes pushed for their inclusion in the deliberative process for deciding which global indicators would be chosen to monitor the 17 SDGs, which resulted in NSOs being designated the voting members of the IAEG-SDGs, while UN agencies, civil society and donor organisations were given the role of observers. NSOs' role is key in this process since, once indicators are set, it is then the NSOs' responsibility to produce the data and statistics to populate these indicators, in order to 'report' on them. In Chart 1, we can see how UN Water has visualised these responsibilities and the flow of data.

We will discuss a very particular part of this data flow—harmonisation—in the next chapter, but here it is important to point out the large number of actors involved in the validation and production of the SDG data[1]. Therefore, the push for more inclusive deliberation was not helped by the problem of the lack of data availability in many countries, which became a development problem in its own right: indeed, the global monitoring of the MDGs had made visible the problem of insufficient data infrastructures for tracking many social, economic and environmental phenomena 'with frequency, timeliness, comparability' (UNSD and FOC, 2014, p. 8). We will discuss the issue of statistical capacity development further in the next chapter, but the anxiety about the sheer number of indicators to be measured at the global level—OWG's original proposal in 2014 included 304 unique indicators in the SDG framework in comparison to the MDGs' 60 indicators—was also

[1] On the country level, apart from the NSO, there is also the larger national statistical system (NSS) that includes ministerial offices that produce statistics and development data. Different governmental officials within the NSS engage with both regional organisations and International Organisations and are required to report data or make sure data is publicly available. It is then International Organisations' responsibility to validate the data with country representatives before sending them on to the United Nations Statistical Division (UNSD) for publication.

acknowledged as a reason to establish a clear process by which proposed indicators would be chosen, including that they have clear methodologies that most countries are producing data to monitor.

The production of indicators in the SDG framework is both a continuation of and an explicit divergence from earlier protocols, networks and institutions used and leveraged to measure economic, social and environmental phenomena on the global level. As discussed in the introduction, a major criticism of the MDGs was that the process by which goals were included in the global agenda was decided from the top without proper consultation with member countries and many bilateral or multilateral development partners. As one member of the global statistics community put it: 'The MDGs were pretty much cooked up by the international agency community' (UN Statistician 4, 2020). This criticism extended to the way that targets and indicators were chosen for inclusion within the MDG framework. In a report on the lessons learned from monitoring the MDGs, the Inter-Agency and Expert Group on the Millennium Development Goal Indicators (IAEG-MDGs) highlighted this as the first weakness ('from statistical but also policy perspective') of the MDGs: 'Targets and indicators were perceived by national statistical systems and other development partners primarily as an international agency driven 'top-down' initiative' (IAEG-MDGs, 2013, p. 3). This was part of the push for more participatory forms of deliberating goals and targets in the SDG framework, which we will discuss further in Chap. 4, as well as the indicators for monitoring progress towards them. Further, there was an underlying argument that the indicators for

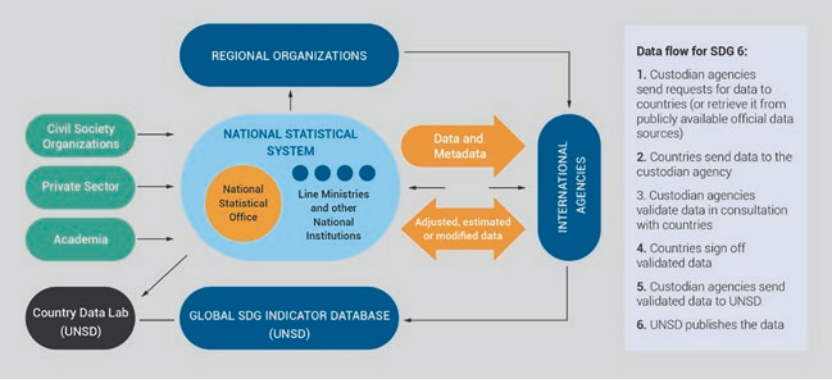

Chart 1 Roles and responsibilities for monitoring and reporting on SDG6 (on water and sanitation), made by UN Water and adapted from the IAEG-SDGs. https://www.sdg6monitoring.org/activities/roles-and-responsibilities/

monitoring the MDGs had been reductionist, and that some of the indicators for measuring societal progress—gross national product (GDP) in particular—needed to be complemented with 'broader measures of progress' within a framework of sustainable development (UNSD, 2013, p. 3). To address these issues, the UN Statistical Commission was called upon to promote 'the science-policy interface through inclusive, evidence-based and transparent scientific assessments, as well as access to reliable, relevant and timely data in areas related to the three dimensions of sustainable development,[2] building on existing mechanisms, as appropriate' (UNGA, 2012, p. 15). The UNSC was established in 1947 to produce and maintain international statistical standards, mirroring and influencing larger trends in international development over more than 70 years of its existence (Ward, 2004). Voting members of the UNSC include representatives of national statistical offices and the statistical offices of International Organisations. Indicators chosen for inclusion in the new, expansive SDG framework had to both incorporate broader definitions of societal progress—including developing indicators whose methodologies and data production were not yet refined—and make use of existing mechanisms 'as appropriate'.

In the same 2013 report on 'lessons learned' mentioned above, the IAEG-MDGs indicated that there were fundamental inconsistencies in these 60 indicators that had been chosen in this 'top-down' manner: 'Some goals, targets and indicators are not well-aligned, and some goals are not adequately addressed by existing indicators' (IAEG-MDGs, 2013, p. 3). Going further, Fukuda-Parr, Yamin and Greenstein (2014, p. 112) argue that it was the indicators and the availability of data that drove which targets were included within the MDG framework,[3] and 'the decision that only targets with agreed-upon indicators and 'robust' data would be included in the goals, with very few exceptions' had a direct impact on derailing the MDGs from the much more expansive Millennium Declaration. However, this emphasis on 'robust' data did not map neatly onto which indicators were chosen, either; the authors found that 'some of the indicators and targets chosen were weakly conceptualized and driven by political considerations as much as measurement ones' (Fukuda-Parr et al., 2014, p. 112). Whether SDG indicators will follow a similar path is an open question—not the least because the IAEG-SDGs have been asked to continue to deliberate on many of the indicators and consider new ones that meet

[2] The three dimensions of sustainable development are the economic, social and environmental.

[3] The MDGs were reviewed three times by the United Nations General Assemblies—in 2005, 2010 and 2013—and at the first review in 2005, the Annual Ministerial Review (AMR) within the UN Economic and Social Council was established to monitor the MDGs annually.

necessary criteria to monitor SDG targets.[4] However, avoiding these inconsistencies was a key impetus for establishing a protocol for reviewing, evaluating and refining indicators.

In March 2013, the UN Statistical Commission (UNSC) established the Friends of the Chair (FOC) group on broader measures of progress as a result of demands made at the 2012 United Nations Conference on Sustainable Development Rio de Janeiro, Brazil (Rio+20) 'to launch a programme of work on broader measures of progress to complement gross domestic product (GDP) in order to better inform policy decisions' (UNSD, 2015a, p. 2). The FOC was explicitly meant to support the intergovernmental process on the post-2015 development agenda—to provide statistical guidance to the Open Working Group (OWG) on Sustainable Development Goals as they discussed, refined and chose the goals and targets to be included within the SDG framework. In this way, the goal to expand beyond the reductionist view of development promoted by GDP was central to the work of the post-2015 agenda as a whole, and to the statistical work of this agenda—as taken on by the UNSC and its associated working groups—in particular. One of the key modes by which FOC—and by extension the global statistical community—provided assistance to the drafting of targets and goals was a set of 29 statistical notes to aid in the OWG's deliberations in March 2014, in which the UNSD and FOC collated and outlined 'main policy issues, potential goals and targets', 'conceptual and methodological tools', 'existing and new indicators' and 'data requirements, challenges and limitations' for 29 varied policy issues (UNSD and FOC, 2014). These notes were meant to 'provide information on the measurement aspects' of those issues discussed by OWG in its first eight sessions (UNSD and FOC, 2014, p. 1).

After extensive deliberation on the targets over the course of 2014, the OWG proposed a list of 304 provisional indicators in communication with FOC for discussion at the 46th UN Statistical Commission in March 2015. In its guidance to the OWG on these provisional indicators in 2015, the FOC used a provisional form of evaluation for each of these initially proposed indicators. The grades for each of the originally proposed indicators were used both by the FOC and in consultation with representatives from national statistics offices from member state countries in order to evaluate each proposed indicator. This provisional form of evaluation gave a

[4] Although the IAEG-SDGs have indeed continued to consider new indicators after its 2020 comprehensive review—in 2021, this included Total Official Support for Sustainable Development (TOSSD), a new indicator for measuring development support under target 17.3—some within the global statistical community say 'this process is basically closed' (UN Statistician 6, 2020).

grade between 'AAA' and 'CCC' to each indicator, where the first letter rates the feasibility of producing data for the indicator, the second rates the suitability of the indicator, and the third letter rates the degree to which the indicator is actually relevant to the target it is meant to measure. The UN Statistical Division, as secretariat of the UNSC, contacted statistical representatives from 70 countries to grade each of the 304 indicators in this way.[5]

At the 46th UNSC in March 2015, the Commission also officially endorsed the establishment of the IAEG-SDGs and tasked the working group to 'fully [develop] a proposal for the indicator framework for the monitoring of the goals and targets of the post-2015 development agenda at the global level, under the leadership of the national statistical offices, in an open and transparent manner' (UNSD, 2015b, p. 1). The group was to consist of 27 representatives of National Statistical Offices (NSOs) and other actors—including international agencies—were to participate only as observers. As with many UN groups, representation was very important, and the IAEG-SDGs were required to ensure 'equitable regional representation and technical expertise and including members of the least developed countries, landlocked developing countries and small island developing States' (p. 1). As might be expected, with the member state NSOs in control of the deliberation of the indicators for inclusion, the relationship between them and the UN agencies and Bretton Woods organisations—as custodian agencies of the indicators—was not clearly defined. As one representative of a member state NSO put it, this relationship was 'a mystery to me actually, even though I was part of it' (National Statistician, 1). The official roles constituted in the IAEG-SDG protocols were not mapped neatly on the actual process for developing and verifying indicators, as this community member remembers it:

> the first time I went to the IAEG-SDG meeting in New York, I couldn't believe how many observers in the form of UN organisations there were. We [the NSOs] were sitting like a small number of people in the midst and everywhere you looked it was like

[5] As a result of this consultation, the proposed indicators were graded in the following way: '50 indicators (16 per cent) were evaluated as feasible, suitable and very relevant (rating AAA) […]. Thirty-nine indicators (13 per cent) received the rating BAA, meaning that those indicators were considered only feasible with strong effort, but suitable and very relevant. Twenty-eight indicators (9 per cent) received the rating BBA, meaning that those indicators were considered only feasible with strong effort, in need for further discussion, but very relevant. Eighty-six indicators (28 per cent) received the rating BBB, meaning that those indicators are considered only feasible with strong effort, in need for further discussion and somewhat relevant. A total of 95 indicators (31 per cent) received the rating CBB, meaning that they were considered difficult even with strong effort, in need for further discussion and somewhat relevant' (UNSD, 2015b, p. 4).

> the sea of UN organisations. But if they had not been there, it would not have worked, because *the stats system is designed for some things, but this system is a lot larger than that*. (National Statistician, 1)

Emphasising the fact that national statistics systems sit within a larger architecture for producing data about global phenomena, she argued that the NSOs could not have produced a global policy monitoring system without the UN agencies. The tension between IOs and member states—which can be found in many different parts of the monitoring process and is a key defining feature of the 2030 Agenda as a whole—is indeed part of what keeps the machinery of global measurement going, as this member makes clear. It also raises the question of how many of the classification and evaluatory decisions made by the IAEG-SDGs are in fact shaped by the path dependency due to many IOs' long histories of producing data about global phenomena.

3.2 The Tier System

The key protocol for evaluating and classifying indicators is the tier system—a tool which was first introduced by the Inter-agency and Expert Group on Gender Statistics (IAEG-GS) in 2012 as a means to evaluate indicators, alongside 'the primary criterion that indicators should address key policy concerns' (UNSD, 2013, p. 3) (Chart 2). The group broke down the tiers in the following way, which map neatly onto the tier system used by the IAEG-SDGs:

March 2015: 46th UNSC in New York	**304 unique indicators discussed (using the AAA-CCC evaluating system); IAEG-SDGs formally endorsed**
June 2015: 1st IAEG-SDGs Meeting in New York	**Tier system provisionally introduced**
March 2016: 47th UNSC in New York	**Tier system formally endorsed**
November 2016: 4th IAEG-SDGs Meeting in Geneva	**230 unique indicators**
March 2017: 48th UNSC in New York	**Formally endorses the IAEG-SDGs' global indicator framework**
July 2017: 71st UNGA	**Adopts the global indicator framework**
March 2020: 51st UNSC in New York	**Tier III indicators eliminated (231 unique indicators, including 130 tier I, 97 tier II, and 4 with multiple tiers)**

Chart 2 Timeline of important moments in the creation and use of the tier system.

Tier 1	Indicators conceptually clear, with an agreed international definition and regularly produced by countries
Tier 2	Indicators conceptually clear, with an agreed international definition, but not yet regularly produced by countries
Tier 3	Indicators for which international standards need still to be developed and not regularly produced by countries

The tier classification process was proposed by the Inter-Agency and Expert Group on the SDGs (IAEG-SDGs) in 2015 and formally agreed upon at the 47th UN Statistical Commission in March 2016 (UNSD, 2016). This classification system was designed as a means to evaluate and refine global indicators for international comparability. At the beginning of the indicator refinement and reclassification process in 2015, 'the largest proportion of indicators was in the so-called tier III category' (UN Statistician, 6). After the IAEG-SDGs' 2020 comprehensive review of the framework—which was approved at the 51st UNSC in March 2020—there were no tier III indicators amongst the 231 unique indicators included in the framework. All tier III indicators had either been eliminated or their methodology had been refined and tested, leading to their reclassification as tier II or tier I indicators. As of December 2020, the 'tier classification contains 130 tier I indicators, 97 tier II indicators and 4 indicators that have multiple tiers (different components of the indicator are classified into different tiers)' (IAEG-SDGs, 2021, p. 2).

Very early on in the IAEG-SDGs' process, however, it became clear that it would be difficult to convert the expansive global agenda of sustainable development into indicators for the SDG monitoring framework. A crucial sticking point for monitoring many targets was that the indicator be 'conceptually clear' in order to be classified as a tier I or tier II indicator. One community member described the difficulty of defining 'sustainable forestry' and 'sustainable agriculture' and 'figuring out' a number in the face of great expectations for the SDGs to be actually transformative and in conversation with other public agencies in her country:

> I think there has been frustrations everywhere because things are complicated and because the people who order the system […] expected it all to be in place very soon. They had no idea how long […] it takes to develop a new statistical thingy. And all across the system, it's like: 'Yeah, we have some data on forestry, but you have asked us for *sustainable* forestry, so now we have to figure out what the criteria would be for that, and then we have to figure out is someone measuring that,' and that's everywhere. I've had discussions with [ministerial] people that were very upset because we had taken the wrong numbers to look at sustainable agriculture. […] I could think of 10 different ways that you could make sustainable agricultural statistics if you wanted to, [but the ministerial representative] didn't want any of these. [Later,] he wrote to me, and said, 'No, I can't put it into numbers what it is that I would like to do, and maybe,' he said, 'we shouldn't have a number for this, then.' [Well,] that's not for us to decide, is it? And being a stats person of course I think it's better *to have something*

> *and then argue about it than having nothing, and everyone thinks that they are talking about the same thing, but they're not.* (National Statistician, 1, our emphasis)

As we will discuss further in Chap. 5, a placeholder number does important work for the global statistical community—allowing for a common language, even if it is not quite the right language yet. The inclusion of 'sustainability' into many of the SDG targets—as a global agenda to support *sustainable development*—has proven very difficult for the IAEG-SDGs and custodian agencies to define statistically, because the definition of sustainability itself is quite open to interpretation.

Further, even if an indicator is conceptually agreed upon, the distinction between tier II and tier III indicators is also about establishing internationally agreed upon standard measurement methodologies and data sources, as discussed further in the next section. In the case of global SDG indicators—as opposed to localised indicators—the onus is on UN custodian agencies and Bretton Woods organisations to do this work of 'reclassifying' the indicators for which they are responsible. This reclassification of indicators requires material and epistemic investment from UN custodian agencies, particularly in pilot testing methodologies, which was felt by some UN agencies as a 'burden':

> The big challenges were to really comply with all the criteria of the reclassification process. And probably the most arduous criterion was having to pilot test the methodology in a regionally representative sample of countries. So, you can immediately imagine that that puts a very big burden on custodian agencies. And compounding that was a situation where some countries were not willing to collaborate for different reasons. Some have their own resource constraints. So even participating in the pilot testing would have entailed some additional resources from their side which they were not able to commit. So, for different reasons there was also, let's say, a reluctance from some countries to participate in the pilot tests, which delayed some of these pilots. (UN Statistician, 6)

Here, we can see how 'reclassifying' indicators for which IOs are responsible is a key mode of asserting their definition of a policy problem on the global policy landscape—one which some countries resisted for both material and 'different reasons'. As we will see below, reclassification is described as a technical process—a matter of leveraging funds to be sure you can run a pilot study in at least one country in each UN region (Africa, Europe and North America, Latin America and the Caribbean, Asia and the Pacific, and Western Asia)—in order to provide sufficient evidence to the IAEG-SDGs to show that an indicator can be feasibly populated with data across the world. However, the reclassification protocol is also one of the key processes by which proposed policy problems can become global public policy problems.

Despite the attempts of the UNSC and its larger community—including the IAEG-SDGs—to promote 'the science-policy interface through inclusive, evidence-based and transparent scientific assessments' (UNGA, 2012, p. 15), there are those in the global statistical community who argue that this promotion of the science-policy interface has been hobbled since the beginning, when statisticians were not invited as official members into the OWG's drafting of the SDG targets. In the original work plan, the drafting of the goals and targets were designated to the politicians, while the indicators were delegated to the statisticians:

> Well, that proved to be maybe a short-sighted, let's say, approach because the result was that the statisticians were not involved in the formulation of the targets, and we have targets that are very wordy, very multidimensional, sometimes requiring many indicators, and the main problem is that they don't specify quantitative thresholds. They use vague terms like increase or something like that. (UN Statistician, 6)

From the perspective of this community, this has complicated the process of monitoring the SDGs, which also fundamentally impairs technical-political accountability: 'The core concern of statisticians with respect to the post-2015 development agenda is the measurability of goals and targets at national and global levels, as a prerequisite for accountability' (UNSD and FOC, 2014, p. 10). At the same time, there has been political pressure on the part of international agencies, member states and civil society to include indicators for their issues in the monitoring framework. As one member of this community described it, 'the SDG indicators were like a big bus [that] some […] people were desperate to get on', as it was clear that 'once [an indicator is] in the framework that will be a powerful measure for countries, for everybody to focus attention of it, use those numbers' (UN Statistician, 3). This push and pull has existed since the beginning of the indicator framework evaluation process: making sure there are *enough* indicators to encapsulate the expansive 2030 Agenda, but also to make sure there are *not too many* indicators that it becomes out of reach to produce monitoring data for many countries with limited statistical capacity.

3.3 The Tier System in Motion: Reclassifying Indicators and Establishing Authoritative Global Public Policy Knowledge

At the 51st United Nations Statistical Commission in March 2020, many representatives of NSOs and statistical offices of International Organisations reiterated and slightly reframed formal congratulations to the IAEG-SDGs for the work entailed

in the 2020 Comprehensive Review of the SDG framework, which included eliminating all tier III indicators. In an illustrative example, the representative of India's statistical office stated:

> we would like to place on record our deep appreciation for the UNSD and [the IAEG-SDGs] towards their efforts for improvement of global indicator framework including tier classification updates, the 2020 comprehensive review of the global indicator framework for the SDGs. Their work on proposed replacement indicators, revisions to existing indicators and proposal for additional indicators are commendable. (Srivastava, 2020, p. 2)

The representative from the United States pointed out that the review process was highly inclusive, and that although the US might have its own ideas about which statistics and data should be used for policymaking, they were committed to using and supporting the global standards decided upon by the working group, calling upon other colleagues in the room to do the same. Coming early in the official statements, this statement from the American representative gave the impression of trying to head off potential contention before it had a chance to be aired. Eliminating all tier III indicators meant that all indicators in the SDG framework had conceptually clear definitions and methodologies, as both tier II and tier I indicators must, but that all countries might not yet be producing the data for the indicators yet, as only tier I indicators do. The IAEG-SDGs had spent all of 2019 and the months before the 2020 UNSC looking over requests for indicator reclassification and refinement, and communicated with custodian agencies that they were required to turn in their supportive materials for this reclassification in time for the working group to review it by January or February of 2020, or else their tier III indicator would be eliminated from the framework—or, to use the language of the UN statistician above, 'kicked off the bus'.

The process of refinement and reclassification of indicators began at the IAEG-SDGs' 3rd meeting in 2016, when the working group first began allowing for indicators to be reclassified and move up or down the ladder of the tier system. Between the 3rd and 4th meetings in April and November of 2016, the first ten indicators were proposed for possible refinement. At the November 2016 meeting, the IAEG-SDGs proposed that the group would 'conduct a review of a set of indicators for re-classification at the Fall physical meeting, once per year'; that agencies would be required to produce their updated information 'at least 1 month before the physical meeting for review by members' and that a 'revised tier classification will only be published once a year following the IAEG-SDG meeting' (IAEG-SDGs, 2016, p. 9). In practice, the working group would hear cases for reclassification at both physical meetings, and sometimes at virtual meetings that happened between the

physical meetings. In the first five years of its existence, the working group physically met twice a year—in addition to six virtual meetings—with the primary objective to reclassify tier III and tier II indicators, to push custodian agencies to test their proposed methodologies and to promote the support for producing data to populate the indicators across the world. After the comprehensive review, the group now only meets once a year, as it is understood that the SDG indicator framework is now—for the most part—complete, and that further refinements or adjustments or reclassification from tier II to tier I will require less work than the scramble that had occupied the group's first five years of existence.

3.4 The Case of Migration: Reclassifying Measures and Policies

Building a case for reclassification is ultimately the responsibility of a custodian agency but must also involve diverse stakeholders. When the IAEG-SDGs announced that their goal was to eliminate the tier III indicators in 2019, the group put out a call to custodian agencies of those indicators that they were to provide the proper methodologies and data sources by the end of 2019, or 'their' indicators would be 'pulled off the table' (National Statistician, 1).

The case of reclassifying the indicator 10.7.2 on migration policies by the UN Department of Economics and Social Affairs (UNDESA), Population Division and International Organization for Migration (IOM) provides a good example of what this process of reclassification looks like in practice (UNDESA and IOM, 2019). As co-custodians of this indicator, UNDESA and IOM were 'required to document, among others, the involvement of governments and national statistical systems in the development of the indicator methodology, and the regional representativeness of the results of pilot studies' (UNDESA and IOM, 2019, p. 11). Therefore, in their case for reclassifying indicator 10.7.2, UNDESA and IOM documented 11 open consultations—taking place from February 2016 to June 2018—which engaged a wide range of stakeholders (almost 300 participants representing governments, International Organisations, NGOs and academics) in the definition of the indicator and the data sources to populate it. They call this process the 'validation of the methodology', and in these consultations, the co-custodians shared their proposed methodologies for feedback.

The methodology validated by this large network of stakeholders included answering questions on an annual survey—the United Nations Inquiry among

Governments on Population and Development (the 'Inquiry')—related to the 30 subsections of the IOM's Migration Governance Framework (MiGOF[6]): the questions asked countries to grade themselves (on a scale of 100) to how well they reached the aspirational framework set out by MiGOF. To produce data required for the reclassification process, UNDESA and IOM launched pilot surveys in 30 countries in order 'to validate and provide additional consistency checks for country responses', and 10 countries responded to the survey with feedback (UNDESA and IOM, 2019, p. 11). In creating an indicator for measuring a country's ability to protect the rights of migrants and promote their well-being, UNDESA and IOM—in conversation with stakeholders—chose the existing data collection mechanism of the 'Inquiry' and limited the scope to the opinions of government entities.

Armed with an explanation of how they chose the MiGOF (the international concept) and the Inquiry (the data source), the documentation of stakeholder involvement and the responses to the pilot survey from the ten geographically representative countries, the UNDESA and IOM put their case forward at the 8th IAEG-SDGs meeting in November 2018. After a brief presentation and discussion at the meeting, the working group decided to grant the request to requalify the indicator to Tier II. UNDESA and IOM were told that in order to be reclassified as a tier I indicator, the data streams for indicator 10.7.2 must be 'established for at least 50 per cent of countries and at least 50 per cent of the population in every SDG region where the indicator is relevant', at which point the co-custodians will submit a new request to the IAEG-SDGs (UNDESA and IOM, 2019, p. 13). The IAEG-SDGs then provided an updated tier classification database to the global statistical community before the annual UNSC in March, seeking official approval from the UNSC. This approval process is largely performative, however, as one member of this community put it:

> the work of the Statistical Commission is really happening […] in those working groups. Normally [the working groups] are tightly scripted—the Statistical Commission gives them a terms of reference and a timeframe and usually also determines the participation and then these groups do their work and then they bring their pieces of work, all the technical work, back to the Statistical Commission, which […] is a four-day meeting, it's a *parliament*. And it's Chief Statisticians, it's not experts on census or civil registration, national accounts or tourist statistics or whatever the topic may be, and then the idea is usually to just have a *high level discussion and wave things through*. (UN Statistician, 3, our emphasis)

[6] https://www.iom.int/sites/g/files/tmzbdl486/files/about-iom/migof_brochure_a4_en.pdf

Therefore, as the 'statistical parliament', the UNSC is crafted as a space for carefully worded congratulations and very few points of substantive discussion about the material of quantification or the policy problems that sit behind them.

3.5 Producing and Eliminating Indicators: The Case of the Tier III Sustainable Tourism Indicator

Of course, when it came down to whether or not 'their' indicators are 'kicked off the bus' of the SDGs, International Organisations and member state representatives vented disagreements about methodologies and data sources. This was the case at the 2020 UNSC, where the IAEG-SDGs presented their comprehensive review for official approval. With the goal of eliminating all tier III indicators[7], the IAEG-SDGs required that all tier III indicators (which included 88 indicators at their peak in 2016) be either reclassified or dropped. In early 2019, the IAEG-SDGs received 251 proposals for changes to the SDG indicator framework. Of these, the working group identified 53 proposals to put towards open consultation, 'including replacements, revisions, additions, deletions and, in a few selected cases, requests for proposals for a group of tier III indicators whose methodological progress has stalled', receiving input from over 600 'individuals/countries/organizations'. For those proposals that included reclassification, this input was in addition to the stakeholder involvement detailed above.

One such tier III indicator whose methodological progress had stalled was indicator 8.9.2, which was worded as 'Proportion of jobs in sustainable tourism industries out of total tourism jobs', and its custodian agency was the World Tourism Organization (UNWTO). The IAEG-SDGs had put forward a proposal for open consultation for replacing this wording with 'Number of employees in tourism industries' to be combined with an additional indicator of 'Energy use by tourism industries' (also a tier III indicator), to address the 'sustainable' component. This was in addition to an existing indicator—designated as tier II and not as threat of elimination—that measures the contribution of tourism to a country's GDP. As the Assistant Director of Statistics Canada argued, without the additional indicator

[7] The review was requested by the UN General Assembly in 2017, which recognised the global SDG monitoring framework 'as a voluntary and country-led instrument that includes the initial set of indicators to be refined annually' and requested that the IAEG-SDGs and the UNSC review the framework comprehensively at the 51st session in 2020 and the 56th session in 2025 (UNGA, 2017, p. 2).

(relating to energy use), 'neither indicator for the target will have anything to do with sustainable tourism' (IAEG-SDGs, 2020).

When reviewing the entire SDG framework, the IAEG-SDGs argued that this new indicator on energy consumption was, indeed, a tier III indicator, and rejected the proposal. Left with one indicator measuring the proportion of GDP produced by tourism and another proposed to measure the number of employees in the tourism industry, the IAEG-SDGs argued that the proportion of GDP and employment actually 'mirror each other, they're almost the same', thus rejected the proposal to include the 'employment numbers' to stand in for 'sustainable tourism' (National Statistician, 1). This elimination of indicator 8.9.2, according to this member of the IAEG-SDGs, was meant to serve as a catalyst in the meeting in order to impress upon UNWTO the importance of better documenting and testing the methodological components of measuring 'sustainable tourism'. However, the temporal frames of the UNWTO—who have been working on an indicator to measure sustainable tourism for 25 years, attempting to grapple with the complexities of defining what sustainability means in the context of tourism—and that of the IAEG-SDGs—facing the deadline of the 2020 Comprehensive Review—clashed, and the indicator was eliminated. At the 2020 UNSC, in the midst of the template congratulations to the IAEG-SDGs, some member states and the UNWTO expressed concern and disappointment that their indicator and policy problem had been 'pulled off the table'. For one member of this group, this was about an 'identity crisis' on the part of International Organisations, which might only have a few indicators in the SDG framework:

> I think for some of […] the UN organisations that have a few indicators might also have just wanted to be able to say that *they have a good influence in the process*. And so, for some of them there is also, even if they don't have a number, they want to have a place there, and then *it becomes more of an identity crisis if we pull things off the table*. (National Statistician, 1)

In the end, as the days of the 2020 UNSC went by, representatives of the Caribbean Community (CARICOM) became increasingly vocal in their disagreement that the indicator should be removed and threatened that they would withhold their approval for the new SDG framework as a result. In this way, for another member of this community, this eruption of disappointment into the normally tranquil space of the UN Statistical Commission was an illustration of what happens for both International Organisations *and* member states when a policy item that is important to them is 'kicked off the bus'. However, it was also an example of how the machinery of the UNSC and its working groups hold together truces—at various levels of fragility—that allow for the infrastructure to continue to function:

> it is rare that the [Statistical] Commission has actually on the floor overruled—I can't now think of an example—because the thing is these things that come to the Commission *are such carefully calibrated compromises*, so if somebody says, 'I will not agree to this package unless you put my tourism indicator in there', then seven others will come up and say, 'well, OK, if you don't agree to this then I don't agree, too—I compromised here and there and there, and then I pull that back', and then the whole thing unravels and then you get to a point—and this is what happened this time again—where people then say, '*OK, we all lose if we let this fall apart*', so grudgingly they agree to that, but it was clear that […] there are a few countries that feel very strongly about—*forget about the whole rest of the SDGs—half of their income is coming from tourism, so then they feel very, very strongly about these things*. (UN Statistician, 3)

In making 'official' decisions on how to measure the world—with what international concepts and with what data sources—the UNSC and its working groups make decisions about what can be counted as global public policy knowledge through these protocols for refinement, addition and reclassification, which are the ultimate responsibility of indicators' custodian agencies. In the next and final section, we will discuss the architecture for producing the data and statistics to populate this global framework.

4 Conclusion

This chapter took a deep dive into the evaluatory and evidentiary practices at the heart of producing the material building block of the SDG framework—the IAEG-SDGs' taxonomy of taxonomies. Through these practices, and with the particular protocols established by the IAEG-SDGs and the larger UN Statistical Commission, member states and International Organisations produce the indicators that work as the essential material underpinnings of the SDG measurement infrastructure. Determining which indicators will materialise or not is a key mode of determining which policy issues get attention, as the SDGs are a bus 'people were desperate to get on' in order to bring attention to their championed issues. The global statistical community has carefully tried to 'gatekeep' the global indicator framework, largely using the argument that if there are too many indicators, then monitoring the entire agenda becomes unreachable for many countries. Yet, this chapter has shown that the SDG governing architecture, with its intricate processes of classifying indicators, has not created an inflexible structure: instead, indicators, seen here as the material representations of social phenomena, are malleable entities, ready to mould, classify and reclassify, upgrade or even eliminate; ironically, elim-

ination here does not mean exclusion. As the case of sustainable tourism showed, countries' pressures and threats can move 'eliminated' indicators back to the negotiating table. Therefore, the production of indicators and the data systems to report on them are not mere 'numbers': they become the material, and absolutely essential, means via which the epistemic infrastructure is built and rebuilt over time and space.

In the next chapter, we will shift to the work that custodian agencies do in producing global public policy knowledge *through* the process of data harmonisation, which is a process of making data comparable. With the crowding of the field of actors producing development data, there has been increased fragmentation of the streams of data produced about economic, political and social phenomena on the national level. Indeed, the problem of statistical capacity in many lower-income countries has arguably been 'baked into' the reasoning for the tier system. The latter and its significance for the production of a global public policy field are the focus of the next chapter.

References

Bandola-Gill, J. (2021). The legitimacy of experts in policy: Navigating technocratic and political accountability in the case of global poverty governance. *Evidence & Policy, 17*(4), 615–633.

Duffy, D. N. (2017). *Evaluation and governing in the 21st century*. Palgrave Macmillan UK.

Espeland, W. N., & Vannebo, B. I. (2007). Accountability, quantification, and law. *Annual Review of Law and Social Science, 3*(1), 21–43.

Fontdevila, C., & Grek, S. (2021). The construction of SDG4: Reconciling democratic imperatives and technocratic expertise in the making of global education data? In S. Grek, C. Maroy, & A. Verger (Eds.), *World yearbook of education 2021: Accountability and datafication in education governance*. Routledge.

Fukuda-Parr, S., Yamin, A. E., & Greenstein, J. (2014). The power of numbers: A critical review of millennium development goal targets for human development and human rights. *Journal of Human Development and Capabilities, 15*(2–3), 105–117.

Inter-Agency and Expert Group on the Millennium Development Goal Indicators (IAEG-MDGs). (2013). *Lessons learned from MDG monitoring from a statistical perspective: Report of the task team on lessons learned from MDG monitoring of the IAEG-MDGs*. United Nations.

Inter-Agency and Expert Group on the Sustainable Development Goal Indicators (IAEG-SDGs). (2015). *First meeting of the inter-agency and expert group on the sustainable development goal indicators New York, 1-2 June 2015. ESA/ST/AC.300/L3*. United Nations.

IAEG-SDGs. (2016). *Plenary 4th meeting of the Inter-agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs) Agenda Item 4. Tier system for the indicators. 17-18 November 2016, Geneva, Switzerland*. United Nations.

IAEG-SDGs. (2020). *Compilation of inputs from the IAEG-SDGs open consultation for the 2020 comprehensive review, 20 september 2019*. United Nations.
IAEG-SDGs. (2021). *Tier classification for global SDG indicators as of 29 March 2021*. United Nations.
Miller, P. (2001). Governing by numbers: Why calculative practices matter. *Social Research, 68*(2), 379–396.
Ordaz, E. (2019). The SDGs indicators: A challenging task for the international statistical community. *Global Policy, 10*(S1), 141–143.
Razavi, S. (2019). Indicators as substitute for policy contestation and accountability? Some reflections on the 2030 agenda from the perspective of gender equality and women's rights. *Global Policy, 10*(S1), 149–152.
Rottenburg, R., & Merry, S. E. (2015). A world of indicators: The making of governmental knowledge through quantification. In R. Rottenburg, S. E. Merry, S.-J. Park, & J. Mugler (Eds.), *The world of indicators* (pp. 1–33). Cambridge University Press.
Shore, C., & Wright, S. (2004). Whose accountability? Governmentality and the auditing of universities. *Parallax, 10*(2), 100–116.
Srivastava, Pravin. (2020). "Statement by Mr. Pravin Srivastava, Chief Statistician of India cum Secretary, Ministry of Statistics and Program Implementation at Fifty first Session of UN Statistical Commission Discussion on Agenda 3(1): Data and indicators for the 2030 Agenda for Sustainable Development." New York: United Nations.
Taylor, D. (2005). Governing through evidence: Participation and power in policy evaluation. *Journal of Social Policy, 34*(4), 601–618.
Ward, M. (2004). *Quantifying the world: UN ideas and statistics*. Indiana University Press.
United Nations Department of Economics and Social Affairs, Population Division and International Organization for Migration (UNDESA and IOM). (2019). *Development, validation and testing of a methodology for SDG indicator 10.7.2 on migration policies*. Technical Paper, available at: https://www.un.org/en/development/desa/population/index.asp
United Nations General Assembly (UNGA). (2017). *Work of the statistical commission pertaining to the 2030 agenda for sustainable development: Resolution adopted by the general assembly on 6 July 2017. A/RES/71/313*. United Nations.
UNGA. (2012). *The future we want: Resolution adopted by the General Assembly on 27 July 2012. A/RES/66/288*. United Nations.
United Nations Statistical Division (UNSD). (2013). *Gender statistics: Report of the Secretary-General. E/CN.3/2013/10.* United Nations.
UNSD. (2015a). *Report of the friends of the chair group on broader measures of progress: Note by the Secretary-General. E/CN.3/2015/2.* United Nations.
UNSD. (2015b). *Technical report by the Bureau of the United Nations Statistical Commission (UNSC) on the process of the development of an indicator framework for the goals and targets of the post-2015 development agenda: Working draft*. United Nations.
UNSD. (2016). *Report on the forty-seventh session (8-11 March 2016). E/2016/24-E/CN.3/2016/34.* United Nations.
UNSD and Friends of the Chair group on broader measures of progress (FOC). (2014). *Compendium of statistical notes for the Open Working Group on Sustainable Development Goals (OWG)*. United Nations.

Harmonising Global Public Policy: Producing Global Standards, Local Data and Statistical Capacity Development

1 Introduction

Following the analysis of the global architecture of indicators for the SDG framework, this chapter will discuss the ways that data production, use and harmonisation have been central to the construction of the materialities underpinning the epistemic infrastructure of the SDGs. For the infrastructure to work smoothly, it is not enough to simply produce the underpinning data and indicators: they also have to be *harmonised* to create a common ground of global knowledge in otherwise fragmented governing spaces. In the language of statistics, harmonisation indicates the 'creation of a desired degree of comparability between statistics of different countries' (Ehling, 2003, p. 17). This construction of the SDGs' epistemic infrastructure has relied on the rise of the use of 'non-traditional sources' of data alongside 'official statistics' to monitor development agendas, and the harmonisation work done by International Organisations and National Statistics Offices to produce these monitoring frameworks. The concept of harmonisation emerges as one of the central ways that International Organisations govern the multiplicity of country-level measures to create universal, global-level metrics. Therefore, it represents a core material element of the epistemic infrastructure: this is the process via which flows of data are produced and the way the infrastructure materially unifies previously disparate parts.

By coordinating the diverse actors (including country governments, civil society, and various experts), IOs aim to achieve the development of *universal mea-*

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_3

sures, allowing for comparison between countries and monitoring of progress across the goals whilst sustaining the country ownership of data and indicators and their active participation in shaping the agenda. Therefore, harmonisation emerges as the central—yet often overlooked—process through which global knowledge and the epistemic infrastructures of global public policy are constructed. Harmonisation is a process through which a variety and diversity of national statistics become translated into one global number.

These universal metrics smooth out profound political, economic and cultural differences between different countries, promote and produce universal policy agendas, and 'create comparability' (Ehling, 2003, p. 17) by promoting competition between countries in various policy arenas. In order to achieve these goals, the process of harmonisation requires constant navigation between country-level measures, grounded in specific political and historical contexts, and global standards, striving for universality and the internationalisation of measurement. In this way, IOs produce the conditions by which countries construct knowledge about their own political, social and economic realities through harmonisation; the latter, in turn, produces and maintains the infrastructure that upholds global public policy, facilitated by commensurability and the production of common goals. We argue that there is no understanding of global governance without an understanding of the processes of harmonisation.

Harmonisation is indeed much more central to the global governance space than standardisation: the two concepts are closely linked but also quite distinct. Global spaces are often assumed to be governed through standards (Ponte et al., 2011)—and even though this is undeniably reflected in the history and practices of the international statistical community, as we will argue in this chapter, it is harmonisation that allows the SDGs to play their central role in creating the global governing space. A concept with a long legacy and multiple meanings, standardisation for the global statistics community encompasses a diverse array of efforts to produce statistical standards in methodologies, interpretation, estimation, dissemination and use of data and statistics. Despite the often-blurred distinction between the two concepts in meetings and official documents, scholars of quantification argue that standardisation is distinct from harmonisation, as the former is the 'complete eradication of difference' while the latter creates a unified field that is not necessarily uniform (Barry, 2001, p. 73).

In the following section, we outline the ways that standardisation and harmonisation have played a key role in producing governable realities, and how the harmonisation of statistics always entwines the scientific and the social (Desrosières, 2000). We then turn to the specifics of how data governance has become particularly complicated with the fragmentation of the global statistics community, as

well as how the tensions that arise with data harmonisation are central to tensions in the 2030 Agenda as a whole. The following section discusses the example of data harmonisation for one SDG indicator—tuberculosis incidence (3.3.2)—to highlight how technical decisions about how to collect, estimate, collate and impute data in order to make global numbers about health are also decisions about how to delineate global public policy. Finally, we discuss how statistical capacity development has become a central stage for the harmonisation of data and the creation of global public policy.

2 Governing Realities: Harmonising and Standardising Data

In the social science and history of quantification and policy, harmonisation and standardisation are specialised modes of creating bound realities: this is particularly important for spatialising and cohering governable entities on national, regional and global scales. James Scott (1998, p. 13) argued that the use of standard spatial measurements and the mapping of land to be governed allowed sovereigns to 'see like a state', which, for example, for the creation of the modern French state hinged on the creation and use of the universal meter. Following Bowker and Star (2000), Timmermans and Epstein (2010, p. 71) define standardisation 'as a process of constructing uniformities across time and space, through the generation of agreed-upon rules'. These standards then 'tend to span more than one community of practice or activity site; they make things work together over distance or heterogeneous metrics; and they are usually backed up by external bodies of some sort, such as professional organizations, manufacturers' associations, or the state' (Timmermans & Epstein, 2010, p. 71). In the realm of producing globally governable entities, Vincanne Adams (2016, p. 24) shows how the disability adjusted life year (DALY) was specifically situated by Bill Gates to become the Lord Chancellor—or the 'one metric to rule them all'—in the context of global health. Adams (2016, p. 29) argues that the DALY, by abstracting 'quality of life and [turning] it into a fiscally meaningful form', standardises ill health globally and makes health universally governable—here, by distinction, converting ill health into an economic concept and an entity to be governed by markets rather than a sovereign. Emerging from a 'crisis of data' in the Global South in the early 1990s, the DALY was meant to provide a universal yardstick for measuring successful or failing health interventions as well as the effectiveness of entire health systems, *creating comparison* between countries that might track health very differently from each other on the national scale.

Harmonisation, however, is distinct from standardisation. Andrew Barry (2001, p. 74) argues that, for the European Union, harmonisation allows for the '"mutual recognition" of national standards' that facilitates unification and a common foundation for international activity in the production of a common 'technological zone' rather than the elimination of difference across borders. The harmonisation that International Organisations take on in the context of the monitoring of the SDGs includes elements of these different formations of governability. Harmonisation, in the context of the European Union, is the process of setting 'the conditions within which a limited degree of standardisation […] is expected to occur' (Barry, 2001, p. 64). Barry shows how creating a unified technological zone was crucial for creating an integrated Europe, and that technological regulation was a key component of this process.

Similar to harmonisation in the context of the European Union, the central United Nations secretariat delegates the creation of standards for harmonising sustainable development across the world. As we discussed in the previous chapter, this is achieved through the production and monitoring of indicators for progression on the SDGs. In using the term 'harmonisation' to describe dealing with difference over a vast array of economic, social, and political institutions, Barry distinguishes it from standardisation by asserting that the former allows for unification while the latter is the 'complete eradication of difference' (2001, p. 73). Rhetorically, at least, this emphasis on harmonisation in the context of the SDGs highlights the fact that actors in the UN space are still trying to unify very different national contexts under the global banner of sustainability, without claiming that they strive to remove all differences.

The technical work of harmonisation is, of course, always political and social. Manfred Ehling (2003, p. 29) refers to 'conflicts of interest' that must be addressed in the process of harmonisation, as 'an abstraction from the different national institutions is needed for the definition of [the] international concept' that is at the centre of harmonisation work. This international concept allows for both input and output harmonisation of data. Ehling breaks down strategies for harmonisation across time and geography into three ideal types, which are useful for distinguishing between different modes of governing data in the context of the monitoring of the SDGs. First, *input harmonisation*—also known as method harmonisation—requires harmonising the tools for data production. This can include, for example, requiring that 'all participating countries use precisely the same survey procedures in an ideal case', like standardising the questions on survey questionnaires (Ehling, 2003, p. 22). Output harmonisation, on the other hand, requires the establishment

of an 'international concept' and involves statistical procedures to convert the 'product of data collection to match that international concept. The second of Ehling's ideal types is *ex-ante output harmonisation*, which—like input harmonisation—uses the design of data production tools as the space for producing harmonised data to capture that international concept but leaves the determination of methods for producing data to individual countries. Finally, in the context of *ex-post output harmonisation*, 'national statistics are subsequently adapted by means of *a conversion procedure* in such a way that comparable statistics can be created' (Ehling, 2003, p. 22, our emphasis). This abstraction and the production of an international concept creates a fundamental tension 'between the quality criteria "international comparability" and "relevance of the (national) statistical concepts"' (Ehling, 2003, p. 22). Because of the differences in national institutions, harmonisation will always create a gap between the nationally relevant concept and the international concept. This gap is wider or narrower depending on how much countries' data production is shaped by the international concept. Alain Desrosières (2000, p. 173) argues that in the context of social statistics—for example, for education, health and poverty—this process of harmonisation is by definition both scientific—'directed at the production of *knowledge*—and social—directed at the production of a *common language* as a foundation for debate on social issues' (his emphasis).

Ideals of standards and harmony do not exist on their own. Like algorithms, techniques of harmonisation for SDG indicators are 'sociomaterial tangles' that are 'composed of collective human practices' (Seaver, 2017, pp. 3, 5). In the context of the SDGs, these techniques of harmonisation—including the production of 'international concepts' to structure data collection and synthesis—require deliberation, the providing of material and evidentiary support, and compromise in mandated spaces for such deliberation, support and compromise within UN agencies, affiliated International Organisations and working groups of the UN Statistical Commission. Most important of all, of course, is the fact that a 'standard or a regulation does not have any natural force or intrinsic momentum. It is an authority which may be obeyed, ignored or opposed' (Barry, 2001, p. 75). The harmonisation and standardisation of data require authority to be taken seriously, and they require country buy-in in order to be implemented on the national or local levels. In the rest of this chapter, we will delve carefully into the 'sociomaterial tangles' of the harmonisation of data for global SDG monitoring, and how these tangles serve work to unify heterogeneous actors under the banner of a global movement for sustainable development.

3 Centrality of Data Production and Harmonisation for the SDGs' Global Public Policy

In order to compare social, economic, political or environmental conditions in two different geographical locations, statisticians harmonise data—by either creating the conditions for producing comparable data or adjusting data after they have been produced—that may have been produced with even slightly different methodologies, including sampling techniques or differently worded answers to questions on household surveys. As many members of the global statistics community made clear in interviews, there is also the importance of harmonising data temporally—being sure to be able to compare contemporary statistics to those that were produced at a different moment in time, when different techniques might have been available—a process that makes statistics 'sustainable'. In the production of the SDGs, harmonisation across space and time happens on many levels, and International Organisations (IOs) and National Statistics Offices (NSOs) engage all (and combinations of all) three of Ehling's ideal modes of harmonisation in the production of comparable data, outlined in the section above. Harmonisation occurs on the global level: IOs 'create comparability' between countries in order to rank performance and identify progress on the SDGs, using both official national data and data produced by donors, civil society and academia. It also happens on the national level, where NSOs are responsible for harmonising data production across different governmental agencies, non-traditional sources of data like geospatial data, and non-governmental and donor-produced data, in order to create a national view of policy problems coherent across different data sources and across time.

3.1 Types of Harmonisation

Outlining the ideal types of harmonisation, an example of (near) *input harmonisation* is UNICEF's Multiple Indicator Cluster Survey (MICS), which is a standardised survey questionnaire that has been used since 1995 to produce comparable data about women and children in 118 countries in the policy arenas of health, education, poverty and more. Data produced by MICS are used both by UN agencies—particularly UNICEF—and countries to monitor progress on poverty, health and education goals. *Ex-post output harmonisation*, however, is of particular use to UN agencies in harmonising data for monitoring the SDGs. This process involves taking nationally produced data—which may not use the same exact survey ques-

tions or identical age ranges in survey sampling, for example—and using statistical tools to adjust situated data that can be compared to others in other contexts. More recently, International Organisations (e.g., the World Bank in their activities on poverty data, e.g., Povcalnet database and the global poverty numbers) started to highlight the value of *ex-ante output harmonisation*. At the level of country dialogue, the IO experts shape the design and collection of the household surveys in ways that then fit the 'global level' requirements.

Statisticians and development data specialists explicitly link the production, harmonisation and use of data to both the production of global agendas and the success of such agendas. From their perspective, there is a danger if data practices do not link up closely enough to the global agenda. Representing Statistics Sweden in the first consultative process in 2015 for the SDG indicators on the part of the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs), Viveka Palm highlighted many difficulties in creating harmonised development data, including that:

> The targets are formulated in words—the data needs to be much more specific. This is the only way to gather data in a harmonised fashion. So, between the policy makers [*sic*] wish and the possible measurement there will often be a gap. The statistics strive to be objective and so they are sometimes hard to interpret in the indicator sense (Measuring the activities in the economy is not equivalent to measure "Sustainable economy" for example.). (IAEG-SDGs, 2015b, p. 382)

Converting political and policy goals—and targets—into statistical measures that can then be 'harmonised' is emphasised by Palm, among other statisticians, as being a key problem faced by members of the global statistical community[1] in bringing the ambitious 2030 Agenda into fruition. There are often 'gaps' between the intention of the targets and what is measurable—both in the sense of what is quantifiable and what data is actually available. Many of the statisticians consulted expressed dismay that they had not been a part of the goal-setting process themselves, in order to help formulate language that would more easily lend goals and targets measurable. However, some scholars have argued that, no matter how participatory the process, the focus on creating and measuring indicators has—like the MDGs before them—simplified a very ambitious agenda into a practice of 'trea-

[1] When we say 'global statistical community', we are referring to a community made up of UN agencies and member states' chief statisticians and statistical staffs, as well as those of Bretton Woods organisations, philanthropic organisations, civil society organisations and public-private partnerships.

suring what we [already] measure', rather than a genuine revolution in what matters in global development (Fukuda-Parr & McNeill, 2019; Yap & Watene, 2019).

Within the UN's 2017 Resolution to adopt the SDG framework, the following roles are set out for IOs and member states around the production, harmonisation and estimation of national data used to monitor the SDGs, where the UN secretariat:

> *Urges* international organizations to base the global review on data produced by national statistical systems and, if specific country data are not available for reliable estimation, to consult with concerned countries to produce and validate modelled estimates before publication, urges that communication and coordination among international organizations be enhanced in order to avoid duplicate reports, ensure consistency of data and reduce response burdens on countries,[2] and urges international organizations to provide the methodologies used to harmonize country data for international comparability and produce estimates through transparent mechanisms. (UNGA, 2017, p. 3)

Built into the SDG monitoring framework, as recognised by the UN, is the responsibility of IOs to harmonise nationally produced data for the purposes of 'international comparability', to make available the means by which they 'produce and validate modelled estimates', and to coordinate with other IOs in order to verify such internationally comparable and sometimes imputed data. In this way, despite the explicit attempts of different UN institutions to guarantee the leadership role of member states in the 2030 Agenda, it is in fact the IOs who are the enforcers of harmonisation. Thus, it is IOs who have the final responsibility to validate the contours of this global public policy—and the processes of harmonisation are of

[2] It is unclear how much this problem has actually been exacerbated by SDG monitoring. One representative of a National Statistics Office argued that this was a key issue that should be addressed by 'harmonisation':

> And even the coordination even between the UN agencies, sometimes it's cost cutting, [the statistical capacity development activities] are repeated between different UN agencies—why are you repeating the same thing? […] So that's why it's also […] coordination and harmonisation that [can save] efforts and money. So, it's not about bringing funds for such a programme for such a region, it's about thinking [bigger] about the focus, about the real capacity development. I also face, we all face within the National Statistical Organisations, each UN agency [works] alone, separately, no kind of coordination between them. For example, we have questionnaires that they sent us to fill in; 90% of the requested data is the same. So, we said, please we're really not, that's so much work on us, we're happy to provide you with the data, but why [not] harmonise? (National Statistician, 3)

key importance to the International Organisations as central to the production of their 'flagship' metrics (such as the International Poverty Line by the World Bank, or the Human Development Index). Consequently, this is the area in which the conflict between IOs and countries and contestation over the global numbers emerge, since the numbers produced via harmonisation might not be the same as the national numbers.

3.2 The 'Data Revolution' and Its Effects on Harmonisation

For many members of the global statistical community, the centrality of data production, harmonisation and use has been sped up by the SDGs. However, it is also the result of the 'data revolution' of proliferated and proliferating digital technologies, whose unequal distribution has exacerbated information inequalities that already existed between the Global South and the Global North. In 2014, UN Secretary-General Ban-Ki Moon commissioned a report on these uneven effects of the data revolution on global development, which was published as *A World That Counts*. The goal of this report was to set out a path towards 'mobilising the data revolution for sustainable development' (IEAG, 2014, p. 2). The authors argue, beyond a common lack of capacity and resources, that too 'often, existing data remain unused because they are released too late or not at all, not well-documented and harmonized, or not available at the level of detail needed for decision-making' (2014, p. 3). As part of a larger movement for evidence-based policymaking, this report reiterated that 'improving data is a development agenda in its own right' and called for a UN-led global public-private partnership—which would become the Global Partnership for Sustainable Development Data (GPSDD)—and an annual global conference on development data—which would become the UN World Data Forum. These communities include both those supporting and harmonising 'official statistics' and those working in parallel streams of 'non-traditional' sources of data—Big Data, private industry, geospatial, academic and so on—with which National Statistics Offices and International Organisations must increasingly contend. Mobilising the data revolution, in this context, means demanding that NSOs and statistical offices of IOs know how to 'filter the wheat from the chaff' on the part of what is useable and what is not (MacFeely, 2019, p. 130).

For statisticians in the UN space, then, the goal is to produce standards for official statistics across both geographic difference and over time. 'Statistics' for this community means both producing data through standardised methodologies and converting these data—as well as alternate streams of data—into official statistics through standardised methodologies. These are standard methodologies that would

allow statisticians to compare poverty rates in two different countries as well as compare poverty rates at two different points in time. With the rise of non-traditional sources of data in SDG monitoring, statisticians see the potential for capturing phenomena that NSOs do not currently capture, which has the potential to provide evidence to vulnerable populations who want to make their perspectives count. However, the rise of non-traditional sources of data creates new problems that statisticians must contend with, as well as new forms of governance both on the part of NSOs and IOs in harmonising data. For some, non-traditional sources of data also raise the question of what has been called the 'sustainability' of statistics, as one statistician put it:

> [One] of the things that I'm also very concerned is the sustainability of statistics itself. Because I mean there are sometimes people that bring data to the table that are from one particular moment in time, because somebody had time, had resources and opportunity to collect the information. And maybe it is a good snapshot of that moment in time, but ultimately that is not what we were interested in. *Because what we are measuring is development*. That means change over time, and my concern is that we have ad hoc collection of information at one point in time and somebody else will decide years later. Inevitably people will divide the newer number by the old number and say something has grown by 10% or discreet increased by X%. But of course, *that statement is only correct if the two methodologies have been identical*. And that is very often not the case with Big Data and ad hoc data collection. So, we are arguing strongly that the national statistical system is the only one you need to institutionalise data collection and perhaps sometimes be a little bit more modest. Collect less but collect it consistently so that after five or 10 years you can really make meaningful assessments of whether you have made progress towards any policy agenda. (UN Statistician, 3)

In order to track change over time, the two data sets must be harmonised—here, in particular, through input harmonisation, that is, the two methodologies for collecting the data at different points in time are 'identical'. This statistician argues that a benefit of harmonisation is that it grants statistics the characteristic of 'sustainability': statistics that are produced now will then continue to have meaning in the future. He contrasts these *sustainable statistics* with 'Big Data and ad hoc data collection', which utilise data production techniques that might not be standardised over time and thus will not be guaranteed to have meaning in the future. From this perspective, measuring development—linked implicitly here to achieving development—*requires* harmonisation and the taking on board of standards.

As non-traditional sources of data and their use have increased in the context of international development since the late 1990s, the development data world has quickly grown and become more and more fragmented—and the work required to

harmonise development data has become more complex in the process. The growing of the field has also led to conflicts over authority of 'official statistics', particularly on the national level, as NSOs have had to engage with data scientists and producers of non-traditional sources of data more and more. One member of the global statistics community argued that the difference between data and statistics is a thorny issue in SDG monitoring, which still requires some effort to overcome, as it is still not clear 'about whose issue' the production of SDG data is:

> I think statisticians feel like this is all their thing, whereas I think some of the people who've been involved, for example, in data science in the private sector see data as something which has a role in industry and in decision-making and in government to some extent that is completely separate from the process of producing official statistics and so I think that debate is to some extent still, I think the different sides, insofar as there are sides, have come to trust each other better, which is good and this is obviously an ongoing debate and it's part of the shift I think within countries about thinking more systemically about their use of data across government and having statistics as part of that, but not the only part of that. (Civil Society, 1)

As an attempt to address these complexities, the 48th UN Statistical Commission, in 2017, called upon the IAEG-SDGs to 'develop guidelines on how custodian agencies and countries can work together to contribute to the data flows necessary to have harmonized statistics' (UNSC, 2017, p. 48). The working group developed guidelines and best practices for harmonising between the national, regional and international levels (IAEG-SDGs, 2018; IAEG-SDGs, 2019). It also produced a series of 'data flow' case studies to understand the 'how an indicator is adjusted, estimated or modelled, and validated by the national statistical system for global reporting', following one indicator from one country to one international agency in each case (IAEG-SDGs, 2017, p. i). These case studies make it clear how the data reporting and harmonisation processes look very different depending on the indicator and the International Organisation responsible for it—and the transparency about how data are converted to match the 'international concept' also varies tremendously. They also make clear the varying role of country-level ministries (and which ministries are) involved in the production of harmonised data. The WHO, for example, states that it has the same relationship with each country for gathering morbidity and mortality data about tuberculosis (through the Ministries of Health and National TB Programs), who reports directly to the WHO at regular points during the year which helps them produce data estimates every year for the country. On the other hand, in the case of 'Indicator 2.1.2 Prevalence of moderate and severe food insecurity', FAO engages with some countries a bit differently, liaising

with NSOs in some cases or with the private company Gallup in others in order to harmonise national data about the indicator.

Conceptualising the harmonisation of data in the context of the SDGs requires dipping into the technical language of statisticians and data scientists, in order to understand the contours of how data is governed for harmonisation. For example, one technique for 'filtering the wheat from the chaff', as MacFeely put it (2019, p. 130), and facilitating input harmonisation for SDG monitoring, is by producing data in such a way that they can be used for multiple purposes. This requires 'interoperability', which is:

> the ability to join-up and merge data without losing meaning (JUDS, 2016). In practice, data is said to be interoperable when it can be easily re-used and processed in different applications, allowing different information systems to work together. Interoperability is a key enabler for the development sector to become more data-driven. (Morales & Orrell, 2018, p. 9)

In this way, interoperability is a technical tool for making data inherently harmonisable. From the perspective of the UN, this requires 'being modest' with data collection (UN Statistician, 3)—setting manageable expectations about what is collectable and programmable to be sustainable.

A particularly contentious space in the production of harmonised data are statistical models and the practice of imputation, the latter of which UNECE defines as a 'procedure for entering a value for a specific data item where the response is missing or unusable' (UNECE, 2000, p. 8). As Brazil's NSO (the Instituto Brasileiro de Geografia e Estatística) asserted in its official assessment of the SDG indicators suggested on 11 August 2015, 'the use of mathematical and/or statistical models to calculate indicators must also be disregarded, since any model is developed according to a given set of assumptions and relatively arbitrary parameters' (IAEG-SDGs, 2015b, p. 4). According to many statisticians, there is some degree of statistical modelling and estimation that is required in producing robust statistical information on countries and their economic and social statistics. In response to IBGE's assertion about the use of statistical models, a representative of FAO argued that:

> Contrary to the implied preoccupation that informs the statement, it is the absence of a proper statistical model in informing an indicator that creates arbitrariness, variability and the impossibility to harmonize measures across countries. The presumption that meaningful indicators could be produced by simple arithmetic computation from primary data collected through censuses or surveys without any statistical treatment is actually a very dangerous one. Models based on sound statistical inference theory are essential, and their use should be broadly promoted, as they are the only

> instrument to ensure a sufficient degree of reliability and comparability of indicators, which should always be seen as estimates of the likely true value of the variable of interest. (IAEG-SDGs, 2015b, p. 39)

Here, this member of the global statistical community is arguing that techniques for adjusting, modelling and imputing data are central not only to the goal of producing harmonisable and meaningful data to monitor global progress towards the SDGs, but also to the production of national statistics in general. The 'dangerous' idea is in fact that data can be made meaningful without 'a proper statistical model'—it is its absence that would make the processing of data 'arbitrary'.

3.3 Neocolonialism? The Creation of 'Parallel Systems'

However, due to a few key actors in the global space—in particular, the Bill & Melinda Gates Foundation (BMGF)—many NSO representatives now understand statistical modelling and imputation as a 'neo-colonial' practice. One member of the global statistics community argued that BMGF has 'largely focused its funding on setting up parallel systems' to national statistics systems, and as a result, in some countries in the Global South, there are 'tens and hundreds of millions of dollars going into parallel systems while the civil registration system is starved of resources, and it's because of this obsession with metrics' (UN Statistician, 13). This creation of parallel systems has real material effects on NSOs and national ministries, and it understandably produces uneasiness with national statisticians on certain technical aspects of producing harmonised data:

> [There's] this tension here which then comes back to haunt us, because I think we've seen a lot of National Statistics Offices in the Global South being disempowered because they lack technically trained staff, so they see the enterprise of estimation and modelling as extremely threatening, see it now as something that's done in Seattle[, Washington] with super computers and people who look like the people who work at Google and Twitter, so it's like that's some kind of rocket science and so they want to cast that as a form of neo-colonialism as opposed to saying, well, hang on a minute, we do need to do modelling, we do need to do estimation. It's even done in statistical agencies in countries like New Zealand and the United States and Canada. *No data system is complete, we're going to miss people. That's just the story of population data and we model, we adjust, there's a whole science around moving from raw data to meaningful and consistent estimates* and that's not just some monopoly which [IHME's Chris] Murray and Gates have, that's something which should be imbued across the entire statistical system of the globe and those skillsets need to be developed in places like Ouagadougou. (UN Statistician, 13, our emphasis)

As discussed in the previous chapter, the capacity to produce, process and use national and sub-national level data is a fulcrum point for the tension between nationally driven development agendas and global agendas like the SDGs. Harmonising data for global monitoring similarly invokes this tension, as IOs enact global public policy through data production, use and harmonisation, processes which we will now outline in the policy areas of health and poverty. This is also an area where the often-contested data practices take place, in order to fill in the missing data (in the cases where the household surveys are missing for a couple of years)—such as nowcasting data or taking country averages.

3.4 Spotlight on Health: The Case of Data Collection on Tuberculosis

Harmonising and monitoring the 27 indicators listed under Goal 3—the health goal—are the responsibility of many UN agencies, including UNICEF, UNAIDS, UNODC, DESA-PopDiv and OECD. Of course, the WHO is at least partly responsible for the large proportion of these (20 of the 27), as well as many health-related indicators situated outside of Goal 3. For each of these indicators, the WHO works with both nationally produced data and its own global estimates of disease burden, which are also modelled on the basis of nationally produced data. Their global estimates are currently produced by the BMGF-funded Institute for Health Metrics and Evaluation (IHME), due to a memorandum of understanding signed between the WHO and IHME in 2018 to partner to produce global estimates (Tichenor & Sridhar, 2020). Since WHO's own team for producing global health estimates, which was led by Colin Mathers, resigned in 2019, IHME 'is now really the only modelling game in town' when it comes to estimates of global disease burden (UN Statistician, 1). Because these estimates can deviate quite widely from numbers reported by ministries of health, and because the WHO is organisationally responsible to its member states, the organisation consults extensively with countries before publishing their global, harmonised health data. To get a sense of what this process of harmonisation looks like for health data, we need to turn to the process of harmonisation for one specific indicator: SDG 3.3.2, 'Tuberculosis incidence per 100,000'.

In the report on case studies of data flows discussed above, the WHO described its method of collecting—as well as estimating—tuberculosis (TB) morbidity and mortality data for all countries in the following way:

> Estimates of the burden of disease caused by TB and measured in terms of incidence (SDG indicator 3.3.2, expressed per 100,000 population per year) and mortality are produced annually by the World Health Organization (WHO), using case notification and death registration information gathered from every country through surveillance systems, special studies (including surveys of the prevalence of disease), mortality surveys, surveys of underreporting of detected TB, in-depth analysis of surveillance and other data, expert opinion and consultation with countries. (IAEG-SDGs, 2017, pp. 22–23)

The authors of the data flows report asked each custodian agency, in a separate category, to describe the 'process by which national data is converted to SDG indicator' (2017, p. 23), where the WHO has inserted its annually updated methodology appendix for estimating TB incidence. In the 2020 version of this methodology paper, Glaziou and his colleagues explain that TB *incidence*[3] (rather than prevalence or mortality[4]) for the global indicator was determined in four ways, depending on the country:

> (*i*) results from TB prevalence surveys (29 countries, 66% of global incidence); (*ii*) notifications in high-income countries adjusted by a standard factor to account for underreporting and underdiagnosis (139 countries, 6% of global incidence) and (*iii*) national inventory studies (8 countries, 17% of global incidence); (*iv*) case notification data combined with expert opinion about case detection gaps (39 countries representing 11% of global incidence in 2019). (Glaziou et al., 2020, p. 1)

The WHO in this way employs both input harmonisation and output harmonisation in its creation of global numbers. For many countries, the organisation also uses multiple methods to create the complete timeline of TB incidence within the years 2000–2019. The ideal mode of producing the SDG indicator for TB, as expressed by the WHO, was through method (*ii*), where countries were recording notifications of new cases of TB through a highly functional health surveillance system.

However, the focus is on TB incidence, rather than mortality or prevalence. This is because the target that the indicator 3.3.2 is meant to measure sets out to 'end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases', among

[3] Within this same document, the authors define incidence as 'the number of new and recurrent (relapse) episodes of TB (all forms) occurring in a given year' within a certain population (Glaziou et al., 2020, p. 24). Prevalence, meanwhile, is defined as 'the number of TB cases (all forms) at the middle of the year'.

[4] Meanwhile, TB mortality data, not included within this particular global indicator for the SDGs, were obtained through national vital registration systems for 123 countries and based on IHME estimates for 21 countries.

other things—a goal acknowledged to be over-ambitious. Tracking incidence 'was estimated with considerable uncertainty in most countries in 2014', but the WHO argues that 'notifications of cases to national authorities provide a good proxy if there is limited under-reporting of detected cases and limited under or over-diagnosis of cases' (WHO, 2021, p. 3). The process by which numbers obtained through these four pathways are 'adjusted by a standard factor to account for underreporting and underdiagnosis' is determined by 'expert opinion', which the WHO has determined in regional workshops 'where expert opinion was systematically elicited following an in-depth analysis of surveillance data' (Glaziou et al., 2020, p. 5). Glaziou and his colleagues set out the limitations of the WHO's mode of estimating TB incidence, including 'a generally small number of interviewed experts; lack of recognition of over-reporting; and others' (Glaziou et al., 2020, p. 6). Further, IHME has also developed an alternative mode of measuring TB burden estimation (Murray et al., 2014), which is 'generally consistent' with WHO's estimates on the global level but varies widely in certain countries. The authors argue that the solution to this problem of 'considerable uncertainty' is statistical capacity development: 'Discrepancies in estimates from different agencies reflect the questionable quality and completeness of the underlying data. Further convergence in estimates will result from improvements in measurements at country level' (Glaziou et al., 2020, pp. 22–23)

Clear in this document on data flows for reporting on TB, too, are complications involved with creating an 'international concept' for monitoring global progress on TB, as well as the ways that scientific and social—to use Desrosières (2000) language—are always intertwined in the creation of quantified governance. Most important is the problem of the co-morbidity of TB and HIV, since TB has become the leading cause of AIDS-related deaths (Pawlowski et al., 2012). Because of the ways that TB and HIV 'act in synergy' to accelerate 'the decline of immunological functions' (Pawlowski et al., 2012, p. 1), indicators to measure TB/HIV have been high on the priority list for organisations invested in combatting TB and HIV, including WHO, UNAIDS, PEPFAR and the Global Fund. Measuring TB prevalence among HIV-positive populations (with the goal of creating a global number for Indicator 3.3.2) in many countries[5] where TB is most prevalent is highly uncertain: it is based on assumptions about HIV prevalence in certain countries and assumptions about the prevalence of TB among HIV-positive populations. The international concept for Indicator 3.3.2 must both encompass and gloss over the syndemic of HIV and TB.

[5] These are in countries that still do not have universal healthcare access and where prevalence data is gathered via survey, as in method (*i*).

Once the WHO's TB programme has its own estimates of TB incidence, annually by August, the organisation then communicates these numbers with countries to verify, and revises its own estimates based on the feedback it receives, before publishing their global numbers every October (WHO, 2021, p. 2). This estimation and verification process predates the SDGs and mirrors other data production and verification procedures for other health programmes within the WHO. The WHO, then, has maintained ministries of health as its main points of contact on the country level rather than national statistics offices, which are the main points of contact for a large part of the global statistics community. As one member of the global statistics community put it:

> WHO is in a position which is slightly different from the other agencies because our constituencies are not the national statistical offices because our constituencies are the ministries of health. So, our interlocutors in country are ministries of health. So, even if we do approach the national statistical office, this usually has to be through the ministries of health. And our governing bodies are ministries of health; they are not the national statistical offices. (UN Statistician, 5)

He went on to assert the fact that collecting and verifying health data is in fact in the 1945 mandate of the WHO, and the SDGs are merely the most recent global agenda to sit upon their already existing data production system. He also asserted that although the WHO encourages countries to create relationships between the ministry of health and the statistical office, they can do nothing to enforce those relationships.

However, from the perspective of UNSD, the harmonisation of data on the country level—and its enforcement—should also be in the remit of custodian agencies. In the context of health, as with other policy sectors, there are multiple ministries on the national level interacting and producing data and statistics that may or may not be comparable within different organisations on the country level. According to one representative of UNSD, it is the responsibility of the NSOs and UN agencies to make sure that this data is harmonised on the country level. This requires 'incentivising' different ministers on the country level, by saying, in the case of health, to these ministry officials:

> 'OK guys, you're doing all those wonderful numbers for health and the Ministry of Health, but have you double checked with the National Statistical Office, are your numbers in sync with the overall population numbers that the Statistical Office manages from the census'. Even stupid little things like age groupings. We have had situations where one ministry has age groups from zero to five, from six to 10 and 11 to 15 and then the next, the other ministry has it grouped from zero to three, from four to seven and so in three-year intervals and so if you then want to conduct any kind of

> study, like how is health and education for instance related to each other, is there any discernible effort if you run an education concern on health behaviour. You can't do it because the two ministries, and that's precisely I think where the National Statistical Office has a role to play to harmonise the frames and, yes, when you work together with others, with the other children then sometimes you have to choose something that is slightly suboptimal for you. (UN Statistician, 3)

Harmonising these frames on the national level allows for different policy arenas to be in communication with each other.

4 Statistical Capacity Development: The Material Production of Development Data

The incomplete, unequal or inconsistent production of official statistics and development data has been framed, particularly since the late 1990s, as its own development problem that requires careful strategy and planning to address. Even before the creation of the Millennium Development Goals and their indicators, certain development plans like the World Bank's Poverty Reduction Strategy Papers, introduced in 1999, and UNICEF's 1990 World Summit for Children put quantified indicators at their centre as a means to measure progress towards development goals and promote adherence to them. The underperformance of many National Statistical Offices (NSO) in the Global South to produce the evidence that these development agendas demanded of them, as well as the 'conflicting donor agendas' (Marrakech Roundtable, 2004, p. 2) that NSOs had to negotiate and that also shifted attention from domestic agendas to global ones, led to a rise in institutional and financial attention to sustainable statistical capacity development—at the time called 'statistical capacity building'.

Support for statistical systems has been categorised—by Rolando Avendano et al. (2021), Shaida Badiee et al. (2017) and others—into two forms: support that takes a 'demand-driven' approach and that which takes a 'supply-driven' approach to development data and official statistics. There is variability amongst and within International Organisations as to which approach they take. The Partnership in Statistics for development in the twenty-first century (PARIS21), for example, firmly supports a 'demand-driven' approach to capacity development, as an organisation whose objective is to advocate for better statistical systems in countries in the Global South for these countries' own development objectives. This is an economic metaphor for indicating that the production of statistics and data should be driven by the demand of countries for evidence to inform their national policymaking, rather than the 'supply-driven' model of data production that has dominated

global governance. Supply-driven statistical capacity development refers to an emphasis on the mere existence of data as evidence of success and not on whether national statistical systems were actually strengthened in the process. Although many UN agencies might not admit to adhering to this philosophy, in practice this latter approach is the dominant one. In the context of the SDGs, the statistical capacity development that UN agencies must facilitate, as custodian agencies of individual indicators, is to provide support for countries to produce data to populate these 231 unique indicators. Some of this funding goes not to statistical systems but instead 'data intermediaries', which include international consultants and global data producers like the Institute for Health Metrics and Evaluation (IHME). In this way, most statistical capacity development within the SDGs is supply-driven, as donors provide funding to help produce the data they need for global monitoring. As the UN Statistical Commission (UNSC) set out roles for the post-2015 development agenda, it:

> stressed the urgent need for investments to enhance national statistical capacity, especially in developing and least developed countries, to measure progress towards the post-2015 development agenda at national, regional and global levels, and enable national statistical offices to play a leading and co-ordinating role in this process. (UNSC, 2016, p. 12)

In this way, the UNSC recognised that one of the key goals of enhancing national statistical capacity was to enable national statistics offices to take leadership over the SDG monitoring framework, and the governing framework by extension. Fundamentally, however, the bulk of the statistical capacity development work that IOs take on facilitates their ability to harmonise data for their purposes of consolidating policy arenas and creating comparability between countries. In other words, much of the work of statistical capacity development in practice in the context of SDG monitoring—despite rhetoric that describes it otherwise—is about producing a terrain for IOs to govern.

In 2016, the United Nations General Assembly commissioned a study by the Joint Inspection Unit to evaluate the work of the UN development system on strengthening national capacities for statistical analysis and data collection, particularly in the support of achieving globally agreed goals, including the MDGs. The authors of the report on the study emphasised how the SDGs would put even more pressure on national statistical offices than the MDGs did, and that although statistics were the *means* to the end of achieving development goals, and not the goal itself, strengthening national statistical systems was critical to making progress on development goals. Thus, achieving the goals would require unprecedented support from the UN system in order to strengthen the production, dissemination

and use of statistics. The authors of the study also identified the great challenges to the success of this support being 'the coordination of activities, the sustainability of the results and the relevance of activities to the priorities of all national stakeholders' (UNJIU, 2016, p. 8).

The landscape for producing statistics and data for national, regional and global development objectives and policymaking has changed in fundamental ways in the twenty-first century. Although attempts to universalise official statistics have been a part of the UN's programme since its inception, there has been a broadening of the statistical community with the 'emergence of quote/unquote, "data science"' in the last decade and a half (UN Statistician, 13). This 'emergence' and rapid rise of data science in international development has produced new epistemic communities, fostered new partnerships and initiatives, and further fragmented the global public policy space. This shifting statistical terrain 'poses serious questions in terms of "what's the role of official statistics" and "what's the particular position of an official statistical agency"', whether that be an NSO or an UN agency or other multilateral or bilateral organisation (UN Statistician, 13). In the evaluation of the World Bank's investment in data for development by its Independent Evaluation Group (IEG), the authors argue that a 'coherent architecture existed for the older generation of partnerships for statistical capacity building, but coherence is missing for the new partnerships involving data innovation' (World Bank, 2018, p. x).

Accompanying the conceptual work of PARIS21 with financial support of statistical capacity development, the World Bank established the Global Trust Fund for Statistical Capacity Building (TFSCB) in 2000 to help countries strengthen the production of their official statistics. At the centre of both of these efforts was the goal of building 'a culture of evidence-based policy making' (Marrakech Roundtable, 2004, p. 2). PARIS21's work varies widely, but in its 2017 mission statement on 'Capacity Development 4.0', Keijzer and Klingebiel (2017, p. 15) argue that 'country ownership' of the development of what they call 'National Statistics Systems' (NSS) is most critical for creating sustainable official statistics production. As a key mode to support this country ownership, PARIS21 helps countries in the Global South develop National Strategies for the Development of Statistics (NSDS) to help create domestic plans for 'evidence-based policy'. This has echoed in other corners of this community, as the authors of that UN the Joint Inspection Unit mentioned above also argued that, with the goal of promoting evidence-based policy, the production and use of statistics must be understood as inseparable: 'It is […] not a case of supporting either production or use of statistics, as the two are intertwined and have a logical linkage' (UNJIU, 2016, p. 9).

Through the institutional, deliberative and financial work of supporting the production of development statistics, International Organisations have actively worked towards harmonising global public policy through the production and use of quantified data. While PARIS21 and TFSCB have had a more holistic conception of developing NSOs and their larger NSSs, many UN agencies have supported statistical capacity development activities that are largely focused on their own policy arenas, like the International Labour Organization's (ILO) activities for bolstering the production of labour statistics or UNESCO's work on bolstering education statistics. The MDGs were fundamental to placing the measurement of indicators at the centre of global public policy, as they explicitly put the responsibility of monitoring and reporting on progress on the shoulders of member states and UN 'custodian agencies' of each indicator. This obligation to monitor and report on poverty, health, education and other indicators also made clear just how many gaps there were in the regular production of social statistics in many countries in the Global South.

The MDGs also brought to the fore a systemic problem in the production of development data and official statistics for global development, as one member of the global statistics community described it:

> in a very aid dependent country, [there are] actually two data systems. There's the data system that the government is painstakingly trying to build [with] inadequate money and not enough people and not enough technology and so on and just to provide that continual feed of information for government decision-making. And then, there's the data system which donors have and fund and is for their own monitoring and evaluation and in line with their own programmes of their preferred surveys for international comparisons and those sort of things, and [often] in a very aid dependent country it's not hard to see how the latter can undermine the former in all sorts of ways, just by diversion of resources, by diversion of people, by warped political incentives, and so on. (Civil Society, 1)

In fact, in their 2005 summary of recommendations, the Inter-Agency and Expert Group on Millennium Development Goals Indicators (IAEG-MDGs) emphasised that 'international agencies should rely more heavily on official statistics produced by national statistical offices for their data needs [and] coordination by donors, bilateral and United Nations agencies in countries should be improved' (IAEG-MDGs, 2005, p. 8). In this way, statistical capacity development represents two processes at once: on the one hand, it is the work of multilateral, bilateral and philanthropic organisations to provide technical and financial support for the statistical system that a country needs for its own programming, and on the other, it is

also the support for the statistical system these International Organisations need for monitoring their own priorities. These two goals for statistical capacity development can—theoretically—align (and sometimes they do), but there are also ample examples of when they do not. In light of these issues, PARIS21 and the UN Statistical Division (UNSD) have attempted to address the problems that these parallel and sometimes contentious double streams of data can produce.

It is a widely held view that the SDGs were a large step towards producing a global agenda that is participatory for all member states. UNSD played a key role in trying to produce a space—both at the annual UN Statistical Commission but also by creating working groups that mandated country participation from various geographic regions. According to one member of the global statistical community:

> It was curious for instance, the role as go-between of the UN Statistical Division, which was somehow overreacting to try to overcome political tensions in terms of having the countries with the perception that it was driven by the UN or by the international system and they didn't want to. So, they went to the other extreme to let all the countries with the national statistical offices […] decide on what to use and which kind of information or indicators should be included. Which in many cases was really an impossible task for the national statistical offices because they were not acquainted with all the domains of development, especially with environmental issues, which were I think supposedly the core of the SDGs. So, we started a process that was painful at the very beginning, but then it went very well in the sense [of starting to work] together to interact as never before. Even from my national experience and then afterwards with the regional and the global experience, I have never witnessed a process that was so really participatory in many aspects. (UN Statistician, 11)

UNSD's introduction of institutional modes for participation included mandating that member states from two different geographic regions co-chair the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs) and that member states have an active role in refining the indicators for each SDG target and goal. As this UN agency representative makes clear, this participatory process was much more involved than those of earlier development agendas and required the active involvement of NSOs—some of whom 'were not acquainted with all the domains of development'—to decide on methodologies and data sources for all aspects of the SDGs.

As the representative for the Samoan and Fijian statistical offices expressed to the IAEG-SDGs in the first SDG indicator consultative process in 2015:

> We believe it is HIGH TIME for all National Statisticians, to see statistical development appear as a development objective in its own right. I hope, madam Chair, as a fellow national statistician, you share in our delight. Having said this, we strongly believe that we require a better indicator, something that builds on the World Bank's

> [Statistical Capacity Index], but which would allow the measurement of 3 core components of national statistical capacity: a. Human capacity (trained, experienced staff to do their job); b. Financial capacity (with Governments providing more than just 'shoe-string' budgets for their NSOs that extend beyond payment of salaries, and actually enables NSOs to do their jobs; and c. Political-institutional capacity, that embraces a culture of evidence informed policy development, planning, monitoring of progress and accounting for results—which requires access to quality and timely statistics. (IAEG-SDGs, 2015a, p. 2)

The SDGs converted statistical capacity development into this 'development objective in its own right'. Perhaps expectedly, consensus about what defines statistical capacity in global governance does not exist. According to one member of the global statistical community (UN Statistician, 7), 'individual organisations are defining capacity and statistical capacity the way they think or their institution mandates to do that. [Consensus] would mean that there are parties who argue about it, [that] there would be a specific discussion, oh, let's define statistical capacity once for all, and then we all concur to that one, which it's like with many other terms, [but] it's not per se happening in this way' for statistical capacity. Because of this lack of consensus, and because of the political power of the SDGs, for many countries the ability to monitor the SDGs themselves becomes a crucial measure of 'statistical capacity'. This has in fact been proposed by some agencies as the way to measure SDG Target 17.19: 'By 2030, build on existing initiatives to develop measurements of progress on sustainable development that complement gross domestic product, and support statistical capacity-building in developing countries'. In this way, International Organisations also direct the statistical capacity through technical and financial assistance that they provide towards supporting the SDG monitoring system itself, with the goal of harmonising and standardising official statistics across the world to make progress universally comparable. For example, the Joint Development Account programme on statistics and data—launched in 2016 with the support of the World Bank, the United Nations Population Fund, the United Nations Development Programme and the European Union, as well as all ten UN agencies of the Account—has the aim of 'strengthening the statistical capacity of developing countries to measure, monitor and report on the implementation of the Sustainable Development Goals, and on progress with regard to their targets and indicators' (UNGA, 2019, p. 15). This recursive quality of measuring statistical capacity development then demands practices of harmonisation on the level of the entire SDG monitoring framework.

To conclude this section, as International Organisations have tried to create it as a development priority in its own right, statistical capacity development has quickly become a 'chaotic' and highly fragmented field: global goals and their monitoring

become more and more central to global progress, in addition to the rise of the use of non-traditional sources of data for monitoring policy goals. As an instrument for influencing national-level policy, as well as a global instrument informed by national priorities, the SDGs' monitoring framework requires increased and sustained financial and technical support that is currently lacking in some countries and geographic regions. For example, according to a representative from the UN Economic Commission for Africa at the 11th Meeting of the IAEG-SDGs in November 2020, there are 52 of the SDGs' 231 unique indicators on which no African country is currently reporting (Ilboudo, 2020). Therefore, like the MDGs before them, although on a different scale and with an explicit attempt at the participatory co-production of global public policy, the SDGs are shaping what is important to measure, and what policies national governments prioritise in the process.

5 Conclusion

This chapter explored the process of harmonisation as central in establishing the building blocks for the epistemic infrastructure of the SDGs. As we argued in the Introduction, data and indicators are the key material manifestations of these structures—and yet, due to the high fragmentation of global governance, IOs continue to be central actors that have to coordinate processes of harmonising these data, as well as support the development of the capacity of national statistical systems so that they are able to continue to produce them.

Indeed, at the start of the twenty-first century, the concept of 'harmonisation' was introduced as a means to address the problem of the outsized power of donor entities—both bilateral funding organisations from the Global North and UN agencies—in directing global public policy. 'Harmonisation' in this context was used to highlight the problem to dual streams of policymaking—those of the 'global', driven by these powerful entities, and those of the 'national' in countries that receive development aid. Statistical capacity development was meant to assert the importance of country-driven (rather than donor-led) development, in the same way that the World Bank's Heavily Indebted Poor Country (HIPC) Initiative was meant to put countries in the driver's seat of their own development plans.

Although the focus in this chapter was specifically on the practice of harmonisation and the development of statistical capacity, in interviews with members of the global statistics community as well as in official documents of key IOs, both 'harmonisation' and statistical capacity development were not used exclusively to describe how data could be made comparable. They were also used to construct pow-

erful and often persuasive narratives of the promotion of country-led development agendas, reaffirming the need for the alignment of donor priorities to countries' priorities, and facilitating an agenda that would be more palatable to the Global South, the participation and approval of which was now more necessary than ever before. It is to the production of these narratives that the next chapter will turn.

References

Adams, V. (2016). Metrics of the global sovereign: Numbers and stories in global health. In V. Adams (Ed.), *Metrics: What counts in global health* (pp. 19–55). Duke University Press.

Avendano, R., Jütting, J., & Kuhm, M. (2021). Counting the invisible: The challenges and opportunities of the SDG indicator framework for statistical capacity development. In S. Chaturvedi, H. Janus, S. Klingebiel, X. Li, A. de Mello e Souza, E. Sidiropoulos, & D. Wehrmann (Eds.), *The Palgrave handbook of development cooperation for achieving the 2030 agenda* (pp. 329–345). Springer International Publishing.

Badiee, S., Klein, T., Appel, D., Mohamedou, E. I., & Swanson, E. (2017). Rethinking donor support for statistical capacity development. In *Development co-operation report 2017: Data for development* (pp. 77–94). OECD.

Barry, A. (2001). *Political machines: Governing a technological society*. Bloomsbury.

Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. (New ed.). MIT Press.

Desrosières, A. (2000). Measurement and its uses: Harmonization and quality in social statistics. *International Statistical Review, 68*(2), 173–187.

Ehling, M. (2003). Harmonising data in official statistics. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison: A European working book for demographic and socio-economic variables* (pp. 17–31). Springer US.

Fukuda-Parr, S., & McNeill, D. (2019). Knowledge and politics in setting and measuring the SDGs: Introduction to special issue. *Global Policy, 10*(S1), 5–15.

Glaziou, P., Dodd, P. J., Dean, A., & Floyd, K. (2020). *Methods used by WHO to estimate the global burden of TB disease*. WHO.

IAEG-MDGs (Inter-Agency and Expert Group on the Millennium Development Goal Indicators). (2005). *Summary of recommendations on capacity-building of the Inter-Agency and Expert Group on Millennium Development Goals Indicators*. United Nations.

IAEG-SDGs (Inter-Agency and Expert Group on the Sustainable Development Goal Indicators). (2015a). *First meeting of the inter-agency and expert group on the sustainable development goal indicators: Statements and related inputs submitted in writing.* ESA/ST/AC.300/6. United Nations.

IAEG-SDGs. (2015b). *Open consultation for members and observers (as of September 2015)*. United Nations.

IAEG-SDGs. (2017). *Data flows from national to regional/global level—Case studies*. United Nations.

IAEG-SDGs. (2018). *Guidelines on Data Flows and Global Data Reporting for Sustainable Development Goals*. United Nations.

IAEG-SDGs. (2019). *Best Practices in Data Flows and Global Data Reporting for the Sustainable Development Goals*. United Nations.

Ilboudo, T. J. (2020). *Capacity development activities in Africa for the 2030 agenda, Economic Commission for Africa*. Presentation given at the 11th Meeting of IAEG-SDGs, virtual.

Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG). 2014. *A World that Counts: Mobilising the Data Revolution for Sustainable Development*. United Nations.

Joined-Up Data Standards project (JUDS). (2016). The Frontiers of data interoperability for sustainable development [online]. Retrieved April 14, 2021, from http://www.publishwhatyoufund.org/wp-content/uploads/2017/11/JUDS_Report_Web_061117.pdf

Keijzer, N., & Klingebiel, S. (2017). *Realising the data revolution for sustainable development: Towards capacity development 4.0*. SSRN Scholarly Paper. ID 2943055. Social Science Research Network.

MacFeely, S. (2019). The big (data) bang: Opportunities and challenges for compiling SDG indicators. *Global Policy, 10*(S1), 121–133.

Marrakech Roundtable on Managing for Development Results. (2004). *Better data for better results. An action plan for improving development statistics (the Marrakech action plan for statistics)*. Presented to the second international roundtable on Managing for Development Results Marrakech, Morocco February 4–5, 2004.

Morales, L., & Orrell, T. (2018). Data interoperability. Global Partnership for Effective Development Cooperation. At https://www.effectivecooperation.org/content/data-interoperability-practitioners-guide-joining-data-development-sector [20.5.2022].

Murray, C. J. L., Ortblad, K. F., Guinovart, C., et al. (2014). Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet, 6736*, 1–66.

Pawlowski, A., Jansson, M., Sköld, M., Rottenberg, M. E., & Källenius, G. (2012). Tuberculosis and HIV co-infection. *PLoS Pathogens, 8*(2), e1002464.

Ponte, S., Gibbon, P., & Vestergaard, J. (2011). *Governing through standards: Origins, drivers and limitations*. Palgrave Macmillan.

Scott, J. (1998). *Seeing like a State*. Yale: Yale University Press.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society, 4*(2), 205395171773810.

Tichenor, M., & Sridhar, D. (2020). Metric partnerships: Global burden of disease estimates within the World Bank, the World Health Organisation and the Institute for Health Metrics and Evaluation. *Wellcome Open Research, 4*, 35.

Timmermans, S., & Epstein, S. (2010). A world of standards but not a standard world: Toward a sociology of standards and standardization. *Annual Review of Sociology, 36*(1), 69–89.

UN Economic Commission for Europe of the United Nations (UNECE). (2000). Glossary of terms on statistical data editing. *Conference of European Statisticians Methodological Material*, Geneva.

United Nations General Assembly (UNGA). (2017). *Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development. A/RES/71/313*. United Nations.

United Nations General Assembly (UNGA). (2019). *Implementation of projects financed from the development account: Eleventh progress report. A/74/85*. United Nations.
United Nations Joint Inspection Unit (UNJIU). (2016). *Evaluation of the contribution of the UN development system to strengthening national capacities for statistical analysis and data collection N1630847*. United Nations.
United Nations Statistical Commission (UNSC). (2016). *Statistical Commission Report on the forty-sixth session (3–6 March 2015) E/2015/24-E/CN.3/2015/40*. United Nations.
UNSC (2017). *Statistical Commission Report on the forty-eighth session (7–10 March 2017) E/2017/24-E/CN.3/2017/35*. United Nations.
World Bank. (2018). *Data for development: An evaluation of World Bank support for data and statistical capacity*. Independent Evaluation Group and World Bank.
World Health Organisation (WHO). (2021). *SDG indicator metadata, harmonized metadata template: Indicator 3.3.2*. United Nations.
Yap, M., & Watene, K. (2019). The Sustainable Development Goals (SDGs) and Indigenous Peoples: Another Missed Opportunity? *Journal of Human Development and Capabilities, 20*(4), 451–467. https://doi.org/10.1080/19452829.2019.1574725.

Scripting the SDGs: The Role of Narratives in Governing by Goals

1 Introduction

Global public policy relies on technocratic governing tools that describe policy priorities using the language and logic of metrics. For some time now, notions such as the efficiency and quality of different policy choices, as well as ideas around regulation and accountability of policy arenas, have been newly reformulated as quantitatively measured entities, made known and available for monitoring and scrutiny by metrics. As the previous chapters have already eloquently shown, comparability of progress, peer pressure and goal-setting have become the primary tools in global governance (Davis et al., 2012).

Nonetheless, despite the central place of metrics in these governing efforts, the translation of goal-setting and numerical thinking into narratives has always been key in adding logic, meaning and sentiment to their cold rationality. Here, we follow Morgan and Wise (2017) to suggest that narrative 'coherence-making' is essential to the fragmented and heterogeneous global governance environment, since narratives are the ordering materialities via which stories are made. These stories 'reveal' and 'unfold' events, meetings, actors and processes over time and place through creating interrelationships: all these material manifestations, written up in executive summaries, PowerPoint presentations, speeches and many other types of texts, uphold the epistemic infrastructure of the SDGs by giving it 'heart and soul'. They are the 'glue' that sticks the infrastructure together, that gives it past, present and a future destination. For what is sustainable development, if it is not a story: a

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_4

story that might have put up divides in the past between the narrators and the narrated, but that now claims that it can also unify and bring together, even momentarily, those who were previously—politically, epistemically or ideologically—conflicted and apart.

Often, the literature is juxtaposing the qualitative formats such as discourses and quantitative forms such as numbers (Porter and Hansen, 2012). In this chapter we challenge this duality (akin to Stone, 2020; Bandola-Gill and Smith, 2021) by exploring different types of narratives in the SDG context. As we will illustrate in the case study of the SDG4, here the narratives might take on a form that is canonically associated with these forms: discursive storytelling of emergency and delay but also stories of the bright utopic future when the goals are realised. However, narratives of the SDGs are shaped beyond words—many of them are also numerical, built around the construction of goals and indicators, whilst others are primarily visual, using the medium of the image to travel and become embedded in the everyday realities of the people they intend to address: not only policymakers but also citizens themselves. Thus, narratives should not be examined as separate from number-making; on the contrary, we see numerical data as key in the construction of narratives about the fabrication of ideal worlds via quantified knowledge. Narratives, through bringing together discursive, numerical and visual elements, become powerful materialities of persuasion and consensus-making. As has been shown by the rich literature in policy and organisational research (McBeth et al., 2012; Radaelli, 1999; Roe, 1994; Czarniawska, 1997), narratives are powerful political tools as they construct and evaluate policy problems and solutions, identify key stakeholders and—perhaps most importantly—offer an interpretative lens on the current political issues which allows to make sense of them and identify potential courses of action. As we will show in this chapter, this narrative-making process in global public policy is both material and multimodal: it facilitates both the stability and the fluidity of the epistemic infrastructure through the interplay between discursive, numerical and visual narratives.

For example, an indicator, in the shortest and simplest form of narrative-making, is not simply a number, but a discursive, numerical and visual story of past, present and future progress. The SDGs, from the start, created the 17 goals and a visual branding that became key in unifying them as a single agenda. The SDG logo, in particular, emphasises the focus on setting such a future agenda, with quantifiable targets, while the colourful wheel highlights both the separation of the goals and their unity (Chart 1).

It is not simply the visual branding that renders text and images key in communicating the message. The combination of the discursive, visual and numerical is vital in the building of individual indicators, too, and in the ways these indicators

Chart 1 The SDG Poster. Disclaimer: The content of this publication has not been approved by the United Nations and does not reflect the views of the United Nations or its officials or Member States. For more information on the SDGs, please visit https://www.un.org/sustainabledevelopment/.

are presented in PowerPoint presentations, sent as email attachments and generally being shared via physical or digital means. The 4.1 SDG indicator[1] is a telling tale of the ways that indicators combine declarative language with the specifics of proportions, time frames and visual imagery to create a narrative of future education (Chart 2).

This chapter will discuss the case of the education SDG (SDG4), as an example of narrative-making in the production of the SDGs. In order to understand how the SDG4 came into being, this chapter uses elements of Narrative Policy Framework (NPF) analysis, applied through an Interpretative Policy Analysis (IPA) lens. Although heavily criticised as taking an almost post-positivistic approach to the study of narratives (Jones & Radaelli, 2015), NPF still offers useful tools to unravel the constituent parts of narrative-making. IPA, on the other hand, is useful here as counter-force to the linearity of the NPF explanation and as a useful sensor to the less discernible, but equally forceful elements of myth-making, central in the bridging of narratives and in creating new global norms (Fontdevila, 2021). The

[1] https://sdg-tracker.org/quality-education

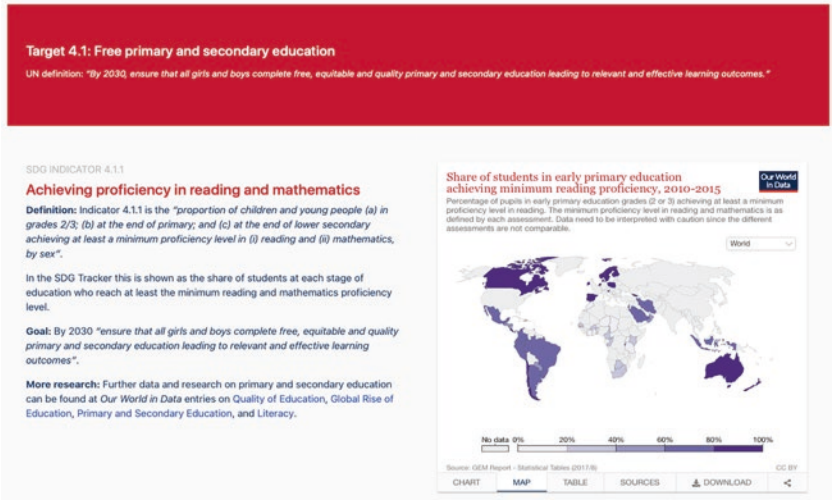

Chart 2 The SDG 4.1 indicator

next section will explain some of the core principles of NPF and the ways that IPA creates useful 'correctives' to such an analysis. We will then move on to the examination of two key texts from the period of preparations towards the production of the SDG4; these are the Muscat Agreement (2014) and the Incheon Declaration (2015). Next, the analysis will turn to a discussion of an education data visualisation, produced for the UNESCO Institute for Statistics (UIS), the key custodian agency of most SDG4 indicators: this is the *Left Behind: Girls education in Africa* visualisation, showing gender education disparities in some of the world's poorest regions. Finally, this chapter will offer a discussion of the role of narratives in shaping new governing norms for global public policy, as well as the close relationships and interdependence of governing narratives with the production of a—mythical—global consensus. As we will see, data visualisations are increasingly used by large International Organisations in their attempts to reach out and make their work appealing and meaningful for donors and countries alike.

2 The Narrative Policy Framework Analysis: Principles, Methods and Criticisms

Narratives are a common form of human cognition; we understand the world around us through stories, and it is stories we use in order to explain our own worlds to others. According to Burke (1973), narratives are 'equipment for living'; they represent stories that have a narrator, a plot and protagonists who are heroes, villains or victims. They often involve a resolution or create counter-narratives and lead to unexpected new beginnings. We grow up with stories, as they are an essential part of the make-up of traditions, of histories, of nations, of our personal trajectories and of artistic creativity itself.

Although many of these observations are fairly commonplace, it is less well-known how narratives are useful in making sense of numbers but also in politicising them in order to achieve specific governing effects. As the visualisation of data over the last few decades has manifestly shown, one of the most powerful communication devices for discussing evidence and numbers are, in fact, stories. There are two main reasons for this: first, narratives become the prime space where numbers are given meaning, and where numbers can be translated into new policy directions (Stone, 2020). Second, according to Espeland, 'analysing the narratives that indicators evoke help us to better understand the effects of quantification' (Espeland, 2015; 61). Paradoxically, one of the major effects of 'narrating indicators', as Espeland argues, is the erasure of narratives in favour of further processes of commensuration:

> Indicators are appealing partly because they simplify complex organisations and processes in order to produce public, authoritative knowledge that makes them appear legible to outsiders. This simplification takes many forms but one way to characterise it is to understand it as the erasure of narratives: the systematic removal of the persons, places and trajectories of the people being evaluated by the indicator and the people doing the evaluation. (Espeland, 2015; 56)

It is this erasure of narratives in the making of numbers that necessitates the need for the construction of new ones. New narratives are necessary to replace 'old' understandings and practices with new re-imaginings and possibilities—the ones that only numbers can create and foster. This is the reason narrative-building is key in quantification; apart from becoming the 'inscribed knowledge' that is 'written down in texts, or represented in pictures and diagrams', numbers narrated in text form have become the only viable way to govern (Freeman & Sturdy, 2014). Historically, the examination of narrative form in social sciences can be traced in

many disciplines: for example, narratives in marketing research were seen as an expert instrument in constructing narrative advertising techniques (Mattila, 2000). In the fields of communication (Morgan et al., 2001) and psychology (Green & Brock, 2000), it was proven that the greater the immersion in a story, the more persuasive its communicative power. Neuroscientists showed the importance of storytelling in individual autobiographical memory and self-conception (see Mar, 2004).

Despite such analyses of narrative-making in a range of disciplines, Jones et al. (2014) in *The Science of Stories* suggest that public policy was slow to acknowledge narrative-making as an essential aspect of the production of knowledge and governing effects, although they do refer to earlier work that focused on narrative analysis from an interpretivist perspective; these were authors like Emery Roe (1994), Fischer and Forester (1993) and Maarten Hajer (1995). Indeed, during the 1990s and in the early 2000s, a number of scholars focused on the role of narratives in shaping policymaking (Berman, 2001; Bleich, 2002; Goldstein & Keohane, 1993; Schmidt & Radaelli, 2004). These authors emphasised the role of ideas in the making of public policy, by showing the ways that policy problems do not always emerge from objective facts and rational interests, but often draw on ideational resources. They, therefore, inserted a significant new element in our understanding of the making of policy stories: that is, the insertion of the argumentative approach in policy analysis. According to Fischer, 'whereas a narrative ties together a story with a beginning, a middle and an end through the device of a plot, an argument is structured around premises designed to logically lead to conclusion' (Fischer, 2003; 181). Thus, both Fischer (2009) and Gottweis (2007) argued that public policy and its narrative forms are not only about *logos* (rational discourse), but also about *pathos* (emotions) and *ethos* (ethics and values).

Finally, the complexity of resources that contribute to narrative-building was further exemplified through Boswell and her colleagues' analysis, who stressed that, apart from the role of ideas, knowledge claims are also key building blocks that narratives draw heavily upon (Boswell et al., 2011); in their words, narratives have to have 'a significant cognitive component which, we argue, creates its own dynamic' (Boswell et al., 2011; 5). According to Boswell and her colleagues,

> As scholars have argued, narratives need to meet certain cognitive criteria. They need to set out causal relations between actions and events (Banerjee 1998; Roe, 1994). We would add that in order to be compelling, they also need to be relatively coherent, consistent with available information, comprehensible, and -in the case of narratives that are scrutinised by researchers- to conform to quite strict criteria of scientific validity. (Boswell et al., 2011; 5)

In order to return to the elements of narrative construction, we will first review the NPF frame of the analysis of story-making. First, NPF proponents stress the need to distinguish between narrative form and content (Jones et al., 2014). In other words, narrative form refers to the structure of a narrative, while narrative content refers to the objects contained therein[2]. In terms of form, narratives use elements that consist of a setting, characters, plot and moral of the story. In more detail, and largely following Jones et al.'s useful outline, here are some of the core elements of the making of narratives:

- First, all narratives need to have a *setting*. The setting is the specific context within which the policy narrative is played out: 'in other words, the setting is the stage, and just like in most plays, people accept the stage as-is without too much thought' (Jones et al., 2014; 6).
- Second, Stone (2002) and Ney (2006) suggest that stories always have *protagonists*: these are characters who are either heroes, villains or victims.
- Third, narratives always have a beginning, a middle and an end—this is the *plot.* Plots are essential in sequencing the story, in inserting causal chains of thought and events. Stone (2002, 2012) has suggested the existence of specific story types, depending on the twists and eventualities that the plot takes: namely, these can be stories of decline, stymied progress, or helplessness and control.
- Finally, narratives usually offer a specific moral; this is the key, 'take- away' message that the story culminates in (such as the ones we are going to examine below). The need for production of more knowledge and hence action is a commonplace moral.

Indeed, as the Introduction has already discussed, the SDGs are addressing some of the most complex, interlinked and compounded global challenges that the

[2] NPF has been a contentious theory and method between these scholars that take a more social constructivist and interpretivist stance and those like Jones et al. who have been attempting very hard to emphasise the social scientific value of the method and the 'objectivity' of the findings that it can offer. According to them, this earlier scholarship of the role of narrative in public policymaking was labelled post-positivist and was, according to them, 'primarily interpretative in the sense that it was highly descriptive, generally rejected scientific standards of hypothesis testing and falsifiability, and thus lacked the clarity to be replicated and allow for generalisation' (Jones et al., 2014; 3). Although a discussion of NPF's epistemology and ontology is beyond the scope of this chapter, our work stresses the interpretivist, social constructivist epistemological conception of NPF, whilst maintaining that it offers a stable and useful empirical set of tools for analysing the construction of policy narratives.

world currently faces. Therefore, the production of narratives that describe these very fluid and often dangerous phenomena is needed all the more; while trying to make sense of these emergencies, narratives also offer some (even momentary) stability and hope. Finally, narratives do not need to be precise; they thus offer added legitimacy to numbers, while masking data gaps and technical inaccuracies. According to Roe (1994; 51) a narrative stabilises 'the assumptions needed for decision making in the face of what is genuinely uncertain and complex. They can be representationally inaccurate—and recognizably so—but still persist, indeed thrive'.

3 Scripting the Narrative of the SDGs

Although quantification has been at the heart of shaping the debate on global goals in both the Millennium and the Sustainable Development Goals, the role of narrative-making has not been investigated in depth. Yet, narratives are key in enveloping and making sense of the data overload; they give numbers meaning and soul. As Shore et al. argue, policies are 'productive, performative and continually contested. A policy finds expression through sequences of events; it creates new social and semantic places, new sets of relations, new political subjects and new webs of meaning' (Shore et al., 2011; 1). It is therefore pertinent to examine the construction of the SDGs, not only as a new measurement agenda comprised of metrics and quantitative data, but also as the construction of a new 'policy world': a new space of political processes, interactions and governing paradigms that becomes consolidated through the use of language and inscription.

Here, we find Barbara Czarniawska's work very useful and especially in relation to her conceptualisation of organisational change and of the vital work that narrative-making does towards producing such change. Czarniawska made substantial contributions to the method of narrative analysis per se, showing its value in understanding organisations:

> Watch how the stories are being made, for example, unfolding how leaders bring together temporality and causality to produce a plausible plot about the necessary course of a change process. Collect stories everywhere: in strategic documents, the boardroom, comics posted on office doors, or the elevator. Provoke storytelling by asking respondents to give their views of what happened first, second…, last, and why. And when moving from field to desk, interpret the stories by asking what people say; analyze the stories, asking how they say it; and deconstruct the stories, asking which perspective are they privileging and which they are silencing. And set narratives together with or against other narratives. It is then time to assemble your own

> story, and theorizing being plotting (Czarniawska & Löfgren, 2013), produce you own theory (Czarniawska, 2014).

One of the most interesting contributions of Czarniawska's work has been her reading of William Tarde and his focus on imitation and fashion: 'Imitation, claims Tarde, is the main mechanism of sociality, the main mode of binding people (and things) to one another' (Czarniawska, 2004, p. 121). Imitation, according to her, is rarely imposed but happens through actors' and organisations' adherence to norms, that is, the things that people 'normally' believe and do. For Czarniawska, fashion is the engine of change: this is the setting of new trends that actors follow through the translation of new discourses in their organisational scripts. Translation in this context does not refer to the precise linguistic act of translating text, but rather understood as 'displacement, drift, invention, mediation, creation of a new link that did not exist before and modifies in part the two agents' (Latour, 1993; 6). According to her, the act of translation is key in the travel of global ideas; she suggests that the new idea, transformed into a buzzword, a model, a presentation or whatever form it may take, is acted upon at national and local policy sites according to need and contextual specificities.

This is the fundamental change that the SDG4 context has generated; building on decades of UN summitry script writing, the SDGs took advantage of a tradition already firmly established: that is, the use of large actor gatherings to explicitly commit and create goal- and target-setting narratives, as the new blueprint for countries to receive, adjust and follow. It is precisely this long tradition of UN work in mobilising public attention and creating new modish ideas for change, that, even in the absence of political consensus, has always sparked optimism in the potential of such large actor congregations—and consequently their declarations and stories—to produce substantial policy shifts. Indeed, despite their focus on initiating change, nothing is ever completely new in these conferences. According to Clark et al. (1998),

> All UN world conferences share similar goals and formats. A central focus of official business at each conference and its preparatory meetings is the creation of a final conference document to be endorsed by state participants. At regional preparatory meetings, governments develop regional positions on specific conference issues. The additional meetings of the Preparatory Committee (PrepComs) are global rather than regional and focus particularly on drafting the conference document. The wording of the final document is invariably the focus of intense politicking among states and between NGOs and states, which continues up to and through the conference. (p. 8)

It is within this background that the narrative of the SDGs was constructed: however, there was also a shift, and one that would mark a sharp difference with all previous UN summitry work. Rather than merely focusing the discussions on the construction of declarations of shared aspiration, typically associated with UN summitry, goal-setting took centre stage. Indicators of progress constituted the bedrock of the new agenda, rather than an additional, technical issue that was only an add-on to the important work of official declarations. Returning to Czarniawska, it is this key operation of goal-setting work that makes the SDGs (and partly their predecessors, the MDGs) and their indicators carry an important discontinuity with previous UN summitry: the production of indicators, enveloped with the well-known declarative language of progress, represented the new modish agenda, able to travel, translate and adjust to national contexts as the new lingua franca of policy innovation and reform.

4 Narratives in and of the SDG4

The section will use two sources of empirical material for the examination of narrative and story-making in the context of the SDG4. It will begin with discourse analysis of two crucial documents in the emergence of the idea and goals of SDG 4, and it will then move on to the analysis of data visualisations, with a specific focus on their examination as storytelling. In presenting both the textual and the visual analyses of narrative and story-making, the aim of this analysis is twofold: first, it will show how through a series of major events and the publication of pivotal texts, such as 'declarations', large global 'agreements' and 'frameworks for action', the work of measurement is inscribed, materialised and made plausible by the production of strong, yet ambiguous, rhetoric of development, equality, democracy, universality and morality. If statistical data is all about possibility, narratives foster *plausibility*; they bring coherence and give sense to the informal and often random and fragmented global governing spaces. We will therefore examine how old and well-established ideas around global development and educational equity and progress are getting new momentum through the use of language that reframes them as large global-setting endeavours. Second, the aim of this analysis is also partly conceptual; the examination of data visualisations will show how the (re-) emergence of the SDG4 narrative requires the complementary work of storytelling to reach out to wider audiences, appeal to local contexts and sentiments and therefore reinforce the narrative in a continuous cycle of bolstering the reach and appeal of the targets themselves.

4.1 The 2014 Muscat Agreement

As indicated in the Introduction, the global governance of education from the start of the century was characterised by the coexistence of multiple, and sometimes, overlapping negotiation processes that were not always harmonious or conflict-free. There were significant power asymmetries and competing expectations in relation to both the decision-making architecture, as well as significant disagreements in relation to the content of these goals and the priorities they placed.

In more detail, since 2000, the global education agenda had been informed by two separate sets of goals; these were, on the one hand, the Education for All (EFA) goals, established in Dakar (WEF, 2000) and, on the other, the MDGs[3]. Importantly, both sets of goals were associated with a specific decision-making architecture and with different communities of practice. Therefore, both agendas emerged in parallel (interestingly, the loci of power were two cities: New York for the EFA and Paris for the MDGs) and from the interaction of different groups of actors, who relied on particular consensus-making scripts; this is significant in relation to the production of narratives, since the EFA negotiations were initially in a power struggle with the MDG education-related goals, only for UNESCO to 'surrender' in the face of a losing battle. While the EFA agenda (and especially, the so-called Dakar goals[4]) was very much the product of a consensus, carefully crafted by the global education community (i.e., specialised circles of experts and transnational bureaucrats with an education background), and reflecting the multiple priorities of education

[3] Education For All (EFA) was a global movement led by United Nations Educational, Scientific and Cultural Organisation (UNESCO), aiming to meet the learning needs of all children, youth and adults by 2015. EFA was adopted by The Dakar Framework in April 2000 at the World Education Forum in Senegal, Africa, with the goal in mind that all children would receive primary education by 2015.[2][3] Not all children receive the education they need or want, therefore this goal was put in place to help those children. UNESCO has been mandated to lead the movement and coordinate the international efforts to reach Education for All. Governments, development agencies, civil society, non-government organisations and the media are but some of the partners working towards reaching these goals.

[4] In 2000, the international community met at the World Education Forum in Dakar, Senegal, an event which drew 1100 participants. The forum took stock of the fact that many countries were far from having reached the goals established at the World Conference on Education for All in 1990. The participants agreed on the Dakar Framework for Action which reaffirmed their commitment to achieving Education for All by the year 2015, and identified six key measurable education goals which aim to meet the learning needs of all children, youth and adults by 2015. In addition, the forum reaffirmed UNESCO's role as the lead organisation with the overall responsibility of coordinating other agencies and organisations in the attempts to achieve these goals.

agencies while allowing civil society to make a meaningful contribution, this was not obviously the case with regard to the MDGs, which viewed education in much more narrow terms (one goal focusing on universal primary education).

Since the negotiation of the SDGs was approached by different agents as an opportunity to put an end to the duality of education agendas, but also as a danger that disagreement in the education policymaking world might mean an exclusion of education from the SDG agenda, the realignment of the EFA agenda with the MDG education 'camp' required crafting a new set of education targets. This meant, in practice, the need to combine the decision-making procedures specific to the EFA architecture with the myriad of negotiation processes set in motion within UN circles as a continuation of the MDGs; thus, the major disagreement was perhaps centred more around the decision-making processes between the two agendas and less about the goals themselves. Nonetheless, it was through a new agreement on the goals that a breakthrough was to be found: this was the Muscat Agreement, signed in May 2014—the document eventually led to its approval at the World Education Forum 2015, with the expectation that it would become an integral part of the global development agenda to be adopted at the UN Summit in New York City in September 2015—that is, the SDGs.

Indeed, the Muscat agreement, signed by a large number of education 'ministers, leading officials of multilateral and bilateral organisations, and senior representatives of civil society and private sector organisations' (p. 1), was the result of the Global Education for All (EFA) meeting in Oman under the auspices of UNESCO's General Conference on 'Education beyond 2015'. It is obvious, even from the very first sentences of this document, that the Agreement and thus the reason for this large gathering of education actors from around the world is not a new development; rather, it is another meeting in the long line of efforts to achieve 'Education for All'. In fact, the document not only does not shy away from its past, but seems to be bolstered by the fact that this appears by now an established and well-trodden path, and one that the EFA 'movement' had established:

> We acknowledge that the worldwide movement for Education for All, initiated in Jomtien in 1990 and reaffirmed in Dakar in 2000, has been the most important commitment in education in recent decades and has helped to drive significant progress in education.

Here, we see that the narrative-building begins through the construction of a shared agenda and a 'movement' that should be not specific to some actors versus others, but that is 'worldwide' and that is marked through important, similar to this one events, in other places and times: in Jomtien in 1990 and Dakar in 2000. As a

result, the text here gathers the pace and progress of past events that have prepared it but also asserts EFA as a significant locus of decision-making in the field. In addition, it also hails the Muscat meeting as a milestone in the line of such agreements and gatherings.

However, the tone quickly shifts and offers an olive branch—neither of the two separate goal-setting 'movements' has achieved their aims:

> Yet we recognise that the Education for All (EFA) agenda and the education-related Millenium Development Goals (MDGs) are unlikely to be achieved by 2015…More than 57 million children and 69 million adolescents still do not have access to effective basic education. In 2011, an estimated 774 million adults, of whom almost two-thirds were women, were illiterate….At least 250 million children are not able to read, write or count…Gender inequality is of particular concern, as only 60% of countries had achieved gender parity at the primary level and 38% at the secondary level by 2011. (p. 1, GEM, 2014)

The use of evidence in narrative-making is a powerful rhetorical tool in creating the necessary epistemic and measurement contexts for launching new decisions and commitments. Startling is also the change of mood here: from the positive and encouraging collective work that has led to this moment (i.e., the Muscat meeting in 2014), numerical evidence is used to show that these efforts still leave a lot to be desired. Therefore, the narration of numbers sets the stage and the mood as one of continuous crisis and emergency: there is urgent need for new action to be taken. Above all, the script is using a certain logic of appropriateness (what is moral and ethical to do) in order to suggest that such evident crisis needs a united policy front, not one riddled with conflict and separation. Such a discourse of consensus-building is core in the production of the narrative in the Muscat Agreement: this is a story about earlier disunity and failure versus a present and future of unity and achievement.

However, the part failure of past efforts does not deter the authors of the text to pace the rate of change; it is precisely the urgency of the situation that further strengthens the commitment to not only achieve the targets previously set but also set new, even more aspirational ones:

> Therefore, we recognise that there is a strong need for a new and forward-looking education agenda that completes unfinished business while going beyond the current goals in terms of depth and scope, as well as to provide people with the understanding, competencies and values they require to address the many challenges that our societies and economies are facing. (p. 2, GEM, 2014)

The Muscat Agreement constructs a narrative that builds on three pillars: first, it clearly spells out that the EFA has been a force of change with a history of over 25 years, the gathering and commitment of key education actors from local, national and international levels; second, the achievement of—at very least—a technical and robust measurement agenda that can offer a fairly concise picture of the levels of educational inequality around the globe and third, the need to unify efforts by both education communities (EFA and the MDGs) in order to have education established as an SDG target in its own right.

Additionally, as the section on 'Vision, principles and scope of the post-2015 education agenda' shows, it works on defining and reaffirming the place of education in—what is slowly emerging as—a global agenda that places sustainable development at its core: it achieves that through outlining the main principles of the group, as well as specifying what the targets for achieving these principles should look like. Interestingly, this is the set of principles that the Muscat participants agreed upon; in summary,

1. 'We reaffirm that education is a fundamental *human right*…';
2. 'The post-2015 education agenda should be *clearly defined, aspirational, transformative, balanced and holistic*, and an integral part of the broader international development framework…*Education must be a stand-alone goal* in a broader post-2015 development agenda and should be framed by a comprehensive overarching goal, with measurable global targets and related indicators…';
3. 'We affirm that *the post-2015 education agenda should be rights-based* and reflect a perspective based on equity and inclusion, with particular attention to gender equality and to overcoming all forms of discrimination in and through education…'
4. We stress that the full realisation of the post-2015 education agenda will require a strong commitment by both governments and donors to allocate *adequate, equitable and efficient financing to education*…accompanied by strengthened *participatory governance*, civil society participation and *accountability mechanisms*… as well as improved *planning, monitoring and reporting mechanisms* and processes' (p. 2, GEM 2014, our emphasis).

We see that there are three primary concerns outlined above; these relate to first, reaffirming the place of education as a human right, therefore connecting closely not only this agreement but also the emergence of the SDG education agenda as a whole with the culture, tradition and institutional identity of UNESCO. This is an important move, as by 2014, multiple other actors, such as the OECD and the World Bank, had also become key education policy-trendsetters globally, and their

perspectives on education did not always coincide with those of UNESCO and the EFA movement. As a result, the Muscat agreement indirectly specifies who the key organisation behind the new post-2015 agenda should be.

Second, the Agreement sets a clear demand for the way forward: education must be a stand-alone goal and not be subsumed by other goals in the SDG agenda. The Muscat Agreement is a key narrative script in—momentarily at least—unifying a vastly conflicted field, that had seen two parallel streams of work emerging globally and often in opposition to each other. Narrative-building starts from three commonplaces for education communities: first, the line of similar events and global meetings; second, the challenges of disagreement and of creating some form of consensus in a really complex and conflicting field and third, the growing crisis of education inequalities. There is a clear message in this narrative that highlights the need to move away from fragmentation towards bringing the two different 'movements' together, in an effort to ensure a singular place of education in the SDG agenda, and not its subsumption within other policy areas and goals.

Where is that the education community should now move to? Goal-setting becomes key narrative practice and takes centre stage in the story, since the signatories appear to universally agree that this should be a technical exercise, focused on pre-defined and well-specified measurement and monitoring agenda, with clear accountability mechanisms and generous funding from donors and governments. The mixing of accountability and financing indirectly connects the two as interdependent. Interestingly, the next section in the agreement moves on to do something quite extraordinary; it sets a number of targets without prespecifying them:

> 'We support "Ensure equitable and inclusive quality education and lifelong learning for all by 2030" as the overarching goal of the post-2015 education agenda.
>
> We further support the translation of this goal into the following global targets, for which minimum global benchmarks and relevant indicators will be identified/ developed:
>
> Target 1: By 2030, at least x% of girls and boys are ready for primary school through participation in quality early childhood care and education…'. (p. 3, GEM 2014)

The list of targets continues with seven targets in total, all of which begin with the time framing of 'by 2030'. They all set specific targets without, however, specifying numerically what the goal should be: in other words, this is a list of 'targets', outlined using language, decontextualised by aspiring them to be applicable globally, yet with no specific numerical inscriptions assigned to them. This practice highlights the 'target-setting' in itself is a narrative-building practice as it creates

'narrative scaffolding' for the policy stories to be told—stories of improvement and mobilisation but also stories of urgency ('by 2030'). What is unique in this case is the fact that this scaffolding is so pervasive that it allows creating numerical narrative, even without specific numbers—just notional percentages of an imagined world 'by 2030'. This is the use of numerical narratives, despite the absence of real numbers quantifying the goals.

Finally, the Agreement ends by explicitly outlining its support to UNESCO to act as the lead organisation for the facilitation of this agenda, in addition to reaffirming the significance of ensuring that the SDG framework has 'a strong education component' (p. 3, GEMA 2014). Although the ambiguity of such non-numbers is startling, what is of interest is the ways numbers still operate as the negotiation instrument for agreeing on a common agenda. Thus, we see the ways that the Muscat Agreement, through its carefully crafted, edited, checked and double-checked script becomes the governing locus where organisational, epistemic and political struggles manage to settle. Its significance is evident, by agreements that succeeded it; first of which was the Incheon Declaration.

4.2 The 2015 Incheon Declaration

The Education 2030 Incheon Declaration was published in the World Education Forum, in Incheon, the Republic of Korea, from 19 to 22 May 2015. According to the document, 'over 1,600 participants from 160 countries' took part; the Forum was organised by UNESCO, 'together with UNICEF, the World Bank UNFPA, UNDP, UN Women and UNHCR' (p. 5, UNESCO, 2016).

The narrative-building in the Incheon Declaration begins from the commonplace of the education emergency that nations are faced with. Nonetheless, it also offers, for the first time, the marrying of the two previous initiatives, in constructing one education goal in the SDG agenda. This is what came to be known as the SDG4—Education 2030 (hence the double-barrelled name):

> 'The world has made some remarkable progress in education since 2000, when the six Education for All (EFA) goals and the Millenium Development Goals (MDGs) were established. Those goals were not, however, reached by the 2015 deadline and continued action is needed to complete the unfinished agenda. With Goal 4 of Transforming our world: the 2030 Agenda for Sustainable Development—'ensure inclusive and equitable quality education and promote lifelong learning opportunities for all' (hereafter referred to as SDG4- Education 2030)—and its associate targets, the world has set a more ambitious universal education agenda for the period from 2015-2030.

> Every effort must be made to guarantee that this time the goal and targets are achieved'. (p. 22, UNESCO, 2016, emphasis in the original)

The document moves on to explain the 'broad consultative process', 'facilitated by UNESCO', which took place to arrive to the SDG4-Education 2030 agenda and targets and further expands on the membership of the decision-making body to now also include a range of actors, such as the OECD; the Global Partnership for Education (GPE); civil society; the teaching profession and the private sector. Therefore, the Incheon Declaration further stabilises the narrative of a universal and aspirational motto of 'education for all' by announcing a single strategy and by adding new, crucial actors the mix of stakeholders agreeing to work together to achieve them; notably, the OECD and the private sector.

Similar to the Muscat agreement, the Declaration is structured around different sections; namely, these discuss 'vision, rationale and principles'; 'the global education goal and its associated seven targets and three means of implementation'; 'governance, monitoring, follow-up and review mechanisms' and finally, 'financing and partnerships' (p. 24, UNESCO, 2016). The focus in this analysis will be centred primarily around the governance and monitoring agenda, as the most relevant in relation to (re-) constructing the narrative of the SDG4.

Although the Declaration begins by referring to the 'old' instruments of establishing principles and values in universal education ('treaties, conventions, agreements and protocols, as well as international instruments, such as recommendations and declarations', p. 31 ibid), it swiftly shifts ground to set a new normal for building global education initiatives. We see a substantial narrative change here towards a transformation to a whole new governing logic, where monitoring, data and accountability are not only important but in fact an indispensable tool for the strategy:

> 'In implementing the new agenda, the focus should be on efficiency, effectiveness and equity of education systems…Furthermore, to ensure quality education and conditions for effective education outcomes, governments should strengthen education systems by instituting and improving appropriate, effective and inclusive governance and accountability mechanisms; quality assurance; education management information systems; transparent and effective financing procedures and mechanisms; and institutional arrangements, as well as ensure that robust, time and accessible data are available'. (p. 32, UNESCO, 2016)

This—importantly—is not only a narrative of the policy contents but rather it offers a new meaning around the governance processes themselves. According to this new narrative, targets should not be open-ended and aspirational declarations

any longer; instead, they have to be 'specific and measurable' and 'country-led' (p. 35)—as such, it proposes both the new heroes of the story (the country government as the key players) but also requires a specific moral to the SDG story, one formulated through precise targets. The Incheon Declaration changes the narrative from previous story-making (e.g., the Muscat Agreement) and suggests that just goal-setting in broad terms will not be enough: instead, there is a need to establish specific targets which will have to be monitored through regular cycles of reporting and accountability: 'this requires establishing intermediate benchmarks (e.g., for 2020 and 2025) through an inclusive process, with full transparency and accountability, engaging all partners so there is country ownership and common understanding'. More explicitly, 'intermediate benchmarks can be set as quantitative goalposts for review of global process vis-à-vis the longer term goals'. Finally, 'intermediate benchmarks are indispensable for addressing the accountability deficit associated with longer-term targets' (p. 35, ibid).

The Incheon Declaration continues the incremental changes pushed by the Muscat Agreement, by offering a measurement-led programme of education governance: the monitoring agenda is not only essential, measurable and country-driven, it also has to be based on a governing architecture with reporting mechanisms at regular intervals through the establishment of intermediate benchmarks. As is commonplace when declaring such substantial shifts in narrative-building, this passage quickly pivots to dramatic language of continued crisis and failure to deliver equitable education for all:

> Despite significant progress since 2000, an estimated 59 million children of primary school age and 65 million adolescents of lower secondary school age…were still out of school in 2013…At least 250 million primary-school-aged children, more than 50% of whom have spent at least four years in school, cannot read, write or count well enough to meet minimum learning standards'. (p. 36, UNESCO, 2016)

The critical turning point that the education emergency has taken requires the drawing up of four different sets of indicators to outline policy priorities and organise the measurement goals: these are specified as global (a small set of globally comparable indicators for all SDGs); thematic (a broader set of globally comparable indicators proposed by the education community); regional and national. Although this differentiation of indicators appeared here as based on levels of government only, it is by now well-documented that eventually it became a qualitative distinction; in other words, much more emphasis has been given to the global indicators (vs. all the other sets) precisely because of the comparability element and the fact that they are part of the SDG framework.

Finally, in terms of 'implementation modalities' (p. 57), national governments are seen as having the 'primary responsibility' or 'regulating standards, improving quality and reducing disparity' (p. 57), following a 'whole of government' approach to education: 'Country-led action will drive change' (p. 60). Interestingly, the document highlights the need for 'regional coordination', too, by suggesting to focus on 'such aspects as data collection and monitoring, including peer reviews among countries; mutual learning and exchange of good practices; policy-making; dialogue and partnerships with all relevant partners; formal meetings and high-level events; advocacy and resource mobilisation; capacity-building; and implementation of joint progress' (p. 61). Thus, the document not only establishes a framework for delivering a measurement agenda; it also creates the expectation that national governments deliver on this agenda and that they do so through peer pressure mechanisms and comparisons with their neighbouring countries and globally.

Therefore, discursively at least, another interesting feature of the new global education narrative in the Incheon Declaration is the repeated emphasis on the need for capacity building in relation to statistical expertise, as well as the 'need for sustained, innovative and well-targeted financing and efficient implementation arrangements'. In fact, the signatories of the Declaration state that the SDG4 targets and policy priorities are explicitly promoted as needing to become part of existing national education policies, plans and processes. It is strongly advised that efforts to realise SDG4 commitments should not result in parallel or separate plans and processes:

> SDG4 policy commitments do not exist outside of existing national policies, planning, management and monitoring processes and mechanisms. Rather, existing country-led systems, processes and mechanisms should be supported or strengthened to ensure better alignment/adaptation with global commitments'. (UNESCO, 2016; 9)

To conclude, it is evident that the SDG4 is not exclusively a performance monitoring agenda. It uses a strong narrative built around it, not only in relation to the need for measurement towards achieving the priorities set (described almost exclusively in the language of different sets of indicators), but also in relation to the new initiative being seen as necessary, ethical, participatory and local. Yet, as we will see in the next section, winning 'hearts and minds' requires more than simply setting up a measurement framework and reporting and accountability mechanisms; it needs a persuasive story.

5 Visualisation as Storytelling in Education

Numerical and discursive narratives are more often than not accompanied by visual narratives. Data storytelling is particularly interesting for the analysis of knowledge production for governing as, instead of concealing the in-built biases and assumptions that all objectivity-making requires, it does precisely the opposite. That is, it works with people's engrained world views as well as learnt visual codes and attempts to mobilise them by pressing towards the making of new political problems and political values.

In this section, we will illustrate this process through which the data is imbued with values and emotions in the *Left Behind*[5]—a data visualisation focusing on girls' education in Africa. It was produced for the UNESCO Institute for Statistics by Function, a data visualisation studio based in Montreal. Its sources primarily draw upon administrative data from UIS. The visual focuses on the gender inequality problem, and in particular the non-participation of African girls in education. As the analysis below will show, although the basis of the *Left Behind* visual is the ranked comparison of African countries and world regions, data and the graphs are simply the setting of the story; the characters, the plot and the moral message are the ones at centre stage (Chart 3).

Chart 3 Front webpage of Left Behind visualisation

[5] https://webarchive.loc.gov/all/20200111114639/; http://uis.unesco.org/apps/visualisations/no-girl-left-behind/

The data visualisation follows very closely the main features of a story; in fact, by using an introduction, as well as specific separate sections, the visual resembles closely the familiar feel and structure of a book. Its title page is very minimal; it offers a title and a subtitle with the background image of a girl reading, while sitting on the ground and leaning back on a wooden structure. More so than the actual image, the colour palette used for the image immediately travels the audience to the dry, hot, dusty African plains—resembling the common depictions of the continent in art and culture.

Against a slightly hazy background (a feature that continues in the whole visualisation), the title fonts are simple, medium-sized and white. There is a certain softness and stillness in the image, as we enter the world of the little girl reading. Despite the crisis in gender equity in education in Africa, the image travels us without any judgements or flashy messages. The title page offers the destination and the focus, while simultaneously creating the sensation of a slow, earthy, hot land where kids still play outside barefoot. The introductory section is structured in a very similar manner: questions ('What would your life be like if you only had 5 years of schooling?'), answers ('For some African girls, this is the most education they can expect, and they are the lucky ones'), and statements of crisis and hope ('Across the region, millions of girls are out of school and many will never set foot in a classroom', 'The world has renewed its promise to the millions of girls who have been left behind'). All the text is presented sentence by sentence as one scrolls through the visualisation, with the background images of girls in classrooms, in the same light creamy, dusky colour hues.

The rest of the visualisation is structured in the format of book chapters, always introduced with a title page (01. The Last Mile, 02. Barriers, 03. Persistence of Illiteracy among Women, 04. Poor school conditions, 05. More Teachers needed, especially women). Each 'chapter' presents relevant data in maps or graph formats. The different pages and graphs are all interactive—they do comparisons of African countries or world regions over time or in ratios. The interactive graphs and maps can be manipulated by viewers through simple movements of the mouse over them. There is nothing extraordinary about these graphs; they follow the common characteristics of contemporary visualisations, following simple lines, laconic explanatory text and modern design.

What is, however, much more interesting when one has a closer look is that all the data charts, maps and graphs are very carefully chosen and put together: some compare selected African countries (depending on the question, these countries are different every time but they are usually low in number). As a result, similar to the image, the data discussed is also fairly minimal, perhaps just a snapshot. Some graphs compare Sub-Saharan Africa with other continents; and others just focus on

simple ratios, between literate and illiterate women. While all data can be accessed by clicking on the black rectangular box at the bottom right of the page, what is striking in every one of these graphs is the careful selection of comparative country or regional data. Although there is clear ranking of countries depending on how well or badly they perform in relation to gender equity, the ranking as a visual, quick and blunt manifestation of best and worst performances is completely abandoned here. There are better and worse country cases (this is the function of any graph and therefore of these graphs, too), but the comparison here only serves as an illustration of the wider political problem of gender inequity—this is further enhanced by the persistent alternating of country comparisons with world comparisons (Chart 4).

Through the mix of data visualisations and other visual elements (pictures, colour palette, the interactivity of the dashboard), *Left Behind* tells a story that balances the need for urgency and intervention and optimistic outlook for the future in which these interventions were realised. Although the main character remains the same (i.e., African girls, women or teachers—as reflected in both the numerical data and pictures), the plot is very carefully crafted in order to move from setting the context outlining the trajectory of change where some challenges have been overcome whilst others took the central stage (0.1 The Last Mile: 'there are good news…but the gender gap persists'), to a discussion of all persisting issues (in 'chapters' 2,3,4) to the relatively uplifting final section on the necessity to have a

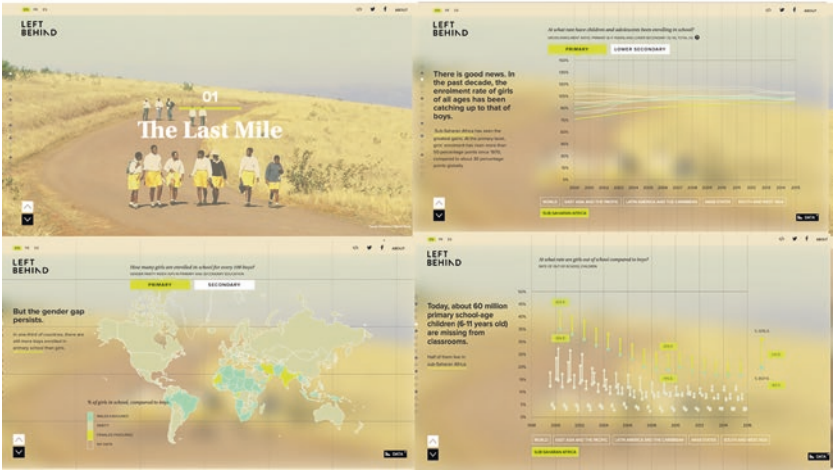

Chart 4 Snapshots of Left Behind visualisation (01)

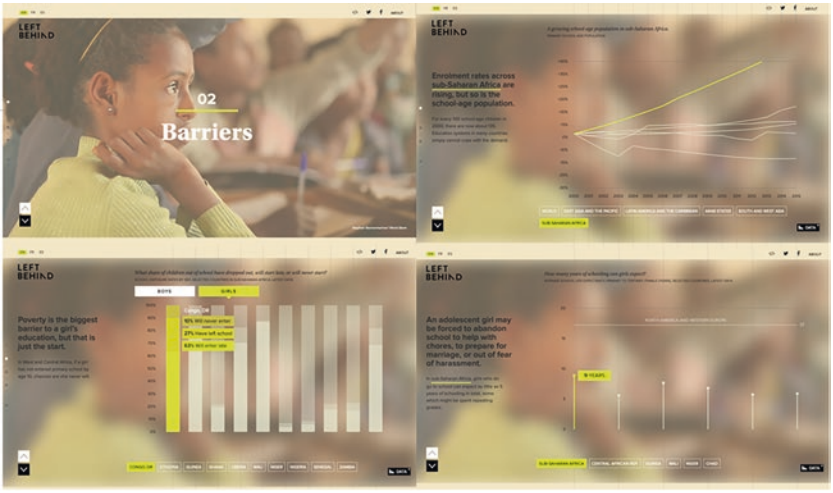

Chart 5 Snapshots of Left Behind visualisation (02)

larger women teacher workforce. Finally, despite what otherwise would have been read as a major inequity crisis, the data visualisation ends the story with nothing less than a 'happy ending': 'The good news is that the international community has not forgotten these girls'. The intention here is for the visual not to paralyse, but fill its viewers with optimism and positive resolve to tackle the problem; and although the text suggests that the SDGs have pledged to decrease inequality, it asks the viewer to also 'have their say' (Chart 5).

This is perhaps the first step in constructing actionable knowledge: enlist one's audience not only to read and understand, but to share their experience of the African girls' education story and mobilise others. Interestingly, the visual does not use any bullet-point language, like most traditional print reports do. While it offers a plethora of interactive information, allowing comparison of performances and progress over time, and although it digests data through some short statements in every page of the analysis, it finishes off with a simple question (Fig. 10): 'What do you think it will take to leave no girl behind?'

This question is at the crux of this chapter's argument: rather than finish off with a definitive memorable statement, or a killer graph, apt for the severity of the issue, *Left Behind* ends with an invitation to the viewer to think for themselves; that is, to weigh the evidence offered and contextualise the issue within their own story-worlds and experiences. Needless to say, this does not mean that

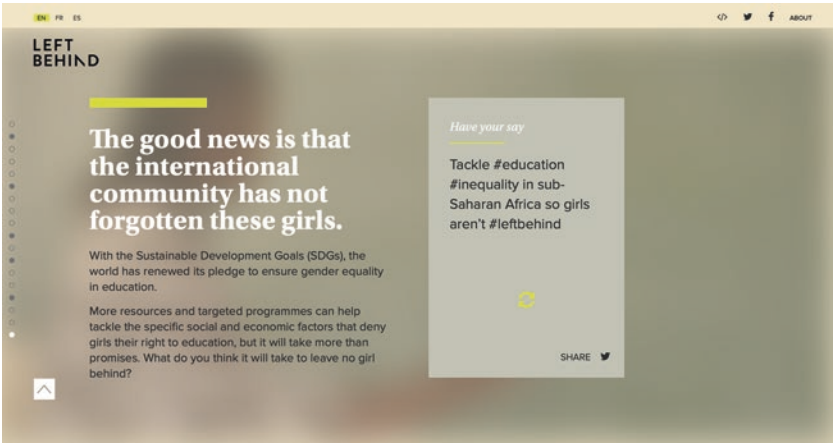

Chart 6 Left Behind last page

careful selection of data and arguments has not taken place here, and that all interpretations and questions are open: quite the contrary. It is precisely because of the meticulous orchestration of text, image and data, as well as the precise crafting of the plot, that this kind of engagement can be invited. In reality, the question is primarily a rhetorical one: these are the multiple worlds that data visualisations fabricate, worlds into which specific and precise policy facts do not matter as much as the interpretive possibilities data (and especially an effective visual data story) can open up (Chart 6).

6 Discussion

This chapter has examined the production of narratives in the field of the global governance of education, especially through the analysis of two major collective declarations as well as the examination of data visualisation as another popular and effective story-making device. We have seen how narratives are used as the material building blocks of the SDG epistemic infrastructure: they work in order to construct sense and stability in situations of fragmentation or increasing complexity. Thus, we see them as vital components of specifying 'who should do what, and how, when and why they should do it in order to address policy dilemmas' (Kaplan, 1986; 770).

It is the potential of narratives to create coherence and consistency of message and structure that makes them particularly necessary as the material underpinnings of the epistemic infrastructure of global governance. As Ricoeur suggests 'the plot or narrative…groups together and integrates into one whole and complete story multiple and scattered events, thereby schematising the intelligible signification attached to the narrative taken as a whole' (Ricoeur, 1984: 10). The intelligibility of events, actors and decisions is of particular significance in global public policy, since the multiplicity of fora, projects and actors renders the field often unknown even to those who are active participants in it. Thus, they do not only create coherence but create logic and, as we saw from the examples above, through the use of 'shocking' numbers of failure they offer compelling and passionate accounts of complex phenomena.

This chapter discussed the ways that stories and narratives in global public policy also depend on creating a crisis discourse; thus, logos (data), pathos (emotion) and ethos (values) are closely intertwined to create calls for unity and action. Data and numbers, therefore, become the engines of story-making: they are not only the valuable resource that allows actors to understand—even feel—the emergency, but through the dominant instrument of goal-setting, metrics are also offered as the vital component of any possible future solution. Here, as suggested earlier, we follow closely Boswell et al.'s conceptual contribution to the study of narratives, which has stressed the cognitive dimension of knowledge claims made. Thus, we see quantified targets as taking centre stage in delineating the nature and scale of the problem (Schneider and Ingram 1993), in constructing causality by constructing arguments that appear comprehensible and convincing and importantly, in appearing themselves as the only viable option for a way forward (Fischer & Forester, 1993); the example of the Muscat Agreement that outlines what the targets and hence policy priorities should be, without specifying numerical figures, is a telling example of this.

Perhaps more importantly, however, we have seen how narratives use goal-setting and numerical targets in order to create bridges and find compromise between otherwise competing and opposing interests and world views. Narratives are inherently formats imposing coherence on complex and messy political realities—and they do so predominantly by selecting ideas and events that could be organised by the plot whilst excluding others. The field of education, and especially the case of the SDG4, is very rich in such a history of education communities being at war with one another, with substantial and enduring differences in relation to both the architecture of governance of the global education policy space and the policy content itself. The scripting of the Muscat agreement is a case in point here: it allowed, after a very long time, the crafting of a narrative that created enough space

and shared targets for both communities to align themselves with, especially under the threat of the exclusion of education from the SDGs as a stand-alone goal. In such a context, goal-setting appears not only as significant instrument for the scripting of the story, but almost as the necessary pre-condition that brings actors together; for if there is one common frame, that is goal-setting as the one globally accepted norm of organising policy work.

On the other hand, the discussion of the *Left behind* visualisation is another eloquent demonstration of the power of numbers to tell a story effectively, using established visual cues and story framings, whilst allowing enough flexibility for the reader to create their own understandings and meanings from the data provided. Thus, it represents the policy narrative's 'multi-interpretability' that provides its appeal to various actors (Hajer, 1995) and enhances the story's persuasiveness and reach. Ultimately, the chapter's main aim was to focus on the materialities and intertwinements of numerical data, discursive text and visual images in order to manifest their vital work in upholding, facilitating and even 'emotionalizing' the epistemic infrastructure. For the infrastructure is not merely a mechanical construction, functioning according to the orders of some distant centre of operations: on the contrary, it is a living macrocosm in need of stories to breathe.

References

Berman, S. (2001). Review article: Ideas, norms, and culture in political analysis. *Comparative Politics, 33*, 231–250.

Bleich, E. (2002). Integrating ideas into policy-making analysis: Frames and race policies in Britain and France. *Comparative Political Studies, 35*, 1054–1076.

Boswell, C., Geddes, A., & Scholten, P. (2011). The role of narratives in migration policy-making: A research framework. *British Journal of Politics and International Relations, 13*(1), 1–11. https://doi.org/10.1111/j.1467-856X.2010.00435.x

Burke, K. (1973). *The philosophy of literary form: Studies in symbolic action*. University of California Press.

Clark, A. M., Friedman, E. J., & Hochstetler, K. (1998). The sovereign limits of global civil society: A comparison of NGO participation in UN world conferences on the environment, human rights, and women. *World Politics, 51*(1), 1–35.

Czarniawska, B. (2004). *Narratives in social science research*. Sage.

Czarniawska, B. (2014). *Social science research from field to desk*. Sage.

Czarniawska, B., & Löfgren, O. (2013). *Coping with excess: How organizations, communities and individuals manage overflow*. Edward Elgar.

Davis, K., Fisher, A., Kingsbury, B., & Merry, S. E. (Eds.). (2012). *Governance by indicators: Global power through quantification and rankings*. Oxford University Press.

Espeland, W. (2015). Narrating numbers'. In R. Rottenburg, S. Merry, S. Y. Park, & J. Mugler (Eds.), *The world of indicators: The making of governmental knowledge though quantification* (pp. 56–75). Cambridge University Press.

Fischer, F. (2003). *Reframing public policy: Discursive politics and deliberative practices*. Oxford University Press.

Fischer, F. (2009). *Democracy and expertise: reorienting policy inquiry*. Oxford University Press.

Fischer, F., & Forester, J. (Eds.). (1993). *The argumentative turn in policy analysis and planning*. Duke University Press.

Fontdevila, C. (2021). Global governance as promise-making: Negotiating and monitoring learning goals in the time of the SDGs. Programa de Doctorate n Sociologia, Departament de Sociologia, Universitat Autonoma de Barcelona.

Freeman, R., & Sturdy, S. (2014). Introduction: Knowledge in policy—embodied, inscribed, enacted'. In R. Freeman & S. Sturdy (Eds.), *Knowledge in policy: Embodied, inscribed, enacted* (pp. 1–17). Policy Press.

Global Education for All Meeting/ UNESCO. (2014). *2014 GEM final statement: The Muscat Agreement*. UNESCO.

Goldstein, J., & Keohane, R. O. (1993). Ideas and foreign policy: An analytical framework. In J. Goldstein & R. O. Keohane (Eds.), *Ideas and foreign policy: Beliefs, institutions, and political change* (pp. 3–30). Cornell University Press.

Gottweis, H. (2007). Rhetoric in policy making: Between logos, ethos and pathos. In F. Fischer, G. Miller, & M. Sidney (Eds.), *Handbook of public policy analysis*. Routledge.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701–721. https://doi.org/10.1037/0022-3514.79.5.701

Hajer, M. A. (1995). *The politics of environmental discourse: Ecological modernization and the policy process*. Oxford University Press.

Incheon Declaration 2015

Jones, M. D., & Radaelli, C. M. (2015). The narrative policy framework: Child or monster? *Critical Policy Studies, 9*(3), 339–355. https://doi.org/10.1080/19460171.2015.1053959

Jones, M. D., Shanahan, E., & McBeth, M. (Eds.). (2014). *The science of stories: Applications of the narrative policy framework in public policy analysis*. Routledge.

Kaplan, T. J. (1986). The narrative structure of policy analysis. *Journal of Policy Analysis and Management, 5*(4), 761–778.

Latour, B. (1993). *We have never been modern*. Harvard University Press.

Mar, R. (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia, 42*(10), 1414–1434. https://doi.org/10.1016/j.neuropsychologia.2003.12.016

Mattila, A. S. (2000). The role of narratives in the advertising of experiential services. *Journal of Service Research, 3*(1), 35–45. https://doi.org/10.1177/109467050031003

Morgan, M., & Wise, N. (2017). Narrative science and narrative knowing. Introduction to special issue on narrative science. *Studies in History and Philosophy of Science, 62*, 1–5. https://doi.org/10.1016/j.shpsa.2017.03.005

Morgan, G., Fischhoff, B., Bostrom, A., & Atman, C. (2001). *Risk communication: A mental models approach*. Cambridge Univ. Press.

Ney, S. (2006). Messy issues, policy conflict and the differentiated polity: Analysing contemporary policy responses to complex, uncertain and transversal policy problems. Vienna, LOS Center for Bergen: Doctoral Dissertation.
Radaelli, C. M. (1999). Harmful tax competition in the EU: Policy narratives and advocacy coalitions. *JCMS. Journal of Common Market Studies, 37*(4), 661–682.
Ricoeur, P. (1984). *Time and narrative*. University of Chicago Press.
Roe, E. (1994). *Narrative policy analysis: Theory and practice*. Duke University Press.
Schmidt, V. A., & Radaelli, C. (2004). Policy change and discourse in Europe: Conceptual and methodological issues. *West European Politics, 27*, 183–210.
Shore, C., Wright, S., & Però, D. (Eds.). (2011). *Policy worlds: Anthropology and the analysis of contemporary power*. Berghahn.
Stone, D. (2002). *Policy paradox: The art of political decision making, revised edition* (3rd ed.). W. W. Norton.
Stone, D. (2012). Transfer and translation of policy. *Policy Studies, 33*(4), 1–17.
UNESCO. (2016). *Incheon declaration and framework for action for the implementation of sustainable development goal 4*. UNESCO.

SDGs and the Politics of Reconciling the Dual Logic of Democracy and Technocracy

1 Introduction

The Sustainable Development Goals (SDGs) mark an important new era in global governance. As previous chapters have already discussed, since the mid-2000s, it has become increasingly clear that global monitoring initiatives must be transformed in order to address grand challenges. Crises such as climate change and the recent pandemic have shown clearly that, although the role of International Organisations (IOs) is key, unless national governments take on the challenge of addressing these emergencies head on, the legitimacy and ultimate success of such efforts is questionable. This has driven to paradigmatic-level change (Best, 2014) in the global governance of shared challenges, where the onus of responsibility for both the decisions and their consequences has moved onto the countries, rather than merely IOs. The design of monitoring programmes and the accountability for their successful introduction and implementation are no longer a matter of IOs' expertise only, but rather are open for negotiations and consensus-building processes for all.

In other words, monitoring regimes have become a subject of both technocratic and democratic logics, and the SDGs—as will be shown in this chapter—are a prime example of this change. The SDGs and their epistemic infrastructure managed to become so extensive precisely because of the types of linkages (or the 'second order' of epistemic infrastructure we outlined in the Introduction) that link more closely country-level decision-makers and global structures of International

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_5

Organisations. Country participation is one of the foundational principles of the SDGs and the priority to 'leave no one behind' explicitly requires all participating actors to be involved. The process of 'democratisation' is not only a matter of equity but also a matter of political buy-in into the infrastructure of measurement within the SDGs.

This, of course, poses a set of fresh challenges for IOs, since the new participatory and country-led paradigm mandates that the SDGs are subject to dual technocratic and democratic legitimacy (Krick, 2018). On the one hand, the monitoring programme is legitimate, based on the technocratic logic of quantification and associated values such as objectivity, expert advice and evidence-based policy-making (Merry, 2016). According to this logic, numbers are powerful as they allow for standardisation and monitoring according to set benchmarks (Hansen & Porter, 2012). The focus on technocratic legitimacy is grounded in the rationalisation (and consequently de-politicisation) of public policy whereby the decision-making is seen as more effective when it is devoid of political pressures (Jasanoff, 2011). On the other hand, the democratic logic surrounding performance exercises sees the value of their embeddedness in national and local politics and highlights the role of social control and transparency over the decision-making process. Furthermore—particularly drawing on STS work on co-production (Jasanoff, 2004) and new modes of knowledge production such as Mode-2 science (Nowotny et al., 2001) and post-normal science (Funtowicz & Ravetz, 1993)—scholars have argued that democratic modes of knowledge production and the opening up the processes of both evidence- and decision-making to the public not only strengthens their legitimacy but also improves their quality as they mobilise multiple viewpoints, values and forms of politics and draw on different knowledge systems, going beyond narrow expert view (Bandola-Gill et al., 2022).

Consequently, the SDGs and their epistemic infrastructures are subject to lines of accountability that might be in contradiction with one another. Focusing on the democratisation of number-based governance is particularly challenging as it poses a challenge to the rules of 'mechanistic objectivity' (Porter, 1995) and might require trade-offs between the professional and scientific standards of statistical reasoning with the political priorities of the usability of the indicators (Bandola-Gill, 2021). The move towards democratisation had important consequences for the processes of developing and implementing the indicators for the SDGs (in particular in contrast to the IO-led MDGs) but also significantly changed the role and position of expertise in the processes of quantification (more on that in Chap. 8—see also Fontdevila & Grek, 2020). This chapter explores this tension in depth by focusing on how technocracy and democracy are navigated within the processes of implementation of the SDGs. This is not only a question of contradictory logics, but

rather promotes a more fundamental discussion of the nature of quantification in contemporary global governance, where the power of numbers is no longer taken for granted, based on their inherent epistemic and political qualities.

2 UN and the 'participatory turn'

As the Millennium Development Goals were progressing, it became increasingly clear that the goals set up would not be achieved (Fukuda-Parr, 2017). Even though the causes of the apparent under-delivery of the MDGs were varied, the commonly discussed reason was its top-down structure and lack of engagement—and consequently buy-in—from the countries involved. Thus, this apparent failure was taken on board in the design of the SDGs and led to a paradigmatic-level change in how the indicator framework for the SDGs was designed and implemented. This change is evident in actors' accounts of the set-up of the monitoring exercise:

> There it was [the MDGs], a very clubby affair. It was basically just us agencies sitting and talking together and all that and very well-meaning of course, but I guess it was a tad elitist in the sense that there are 20 people in a room versus 200. […] So, just that type of dialogue and all that we didn't have before the SDGs, and also dialogue with countries. At first, the countries were very much, naturally—they were very annoyed at the international agencies being in the front seat and them being in the back seat. This is a country-led process and it was completely flipped and then there was the discomfort with that also, because how can we have you measure something that you are judging your own progress by; it's like you grading your own paper. But I think, so the entente has been reached and there is, I think the statistical world will be better for it. (World Bank, 15)

This quote clearly illustrates two key tensions embedded in this new paradigm of engagement in the global monitoring system—on the one hand, it breaks with the Western-centric, elitist view of the development and promotes more equal participation by the countries who are most affected by these systems. On the other hand, such 'opening up' risks methodological challenges, as it necessarily involves countries in the politics of measurement in a much more direct way.

The participatory turn of the UN monitoring system occurred not merely at the level of procedural 'behind the scenes' politics, but rather, it was embedded in the key document establishing the SDGs. The flagship document of the Rio Conference (see: Chap. 1)—*The Future We Want*—was at its core a political declaration of inclusivity of the different voices into the governance of the SDGs. For example:

> We reaffirm the key role of all levels of government and legislative bodies in promoting sustainable development. We further acknowledge efforts and progress made at the local and sub-national levels and recognize the important role that such authorities and communities can play in implementing sustainable development, including by engaging citizens and stakeholders and providing them with relevant information, as appropriate, on the three dimensions of sustainable development. We further acknowledge the importance of involving all relevant decision-makers in the planning and implementation of sustainable development policies. (UN General Assembly, 2012, p. 8)

As evident in this quotation, the inclusion of not only the policymakers but also a range of other stakeholders (such as the civil society and national representatives) was seen as necessary for the SDGs' success. This—importantly—was a paradigm set for all the countries, and not only rich funders from the North. *The Future We Want* (UN General Assembly, 2012) explicitly discusses the involvement of developing countries as equal and necessary participants in sustainable development governance. As indicated in the following:

> We reaffirm the importance of broadening and strengthening the participation of developing countries in international economic decision-making and norm-setting, and in this regard take note of recent important decisions on reform of the governance structures, quotas and voting rights of the Bretton Woods institutions, better reflecting current realities and enhancing the voice and participation of developing countries, and reiterate the importance of the reform of the governance of those institutions in. (UN General Assembly, 2012, p. 19)

These political declarations went even further in 'Transforming Our World' (UN General Assembly, 2015), the cornerstone document, establishing the SDGs as a political programme. The SDGs from the outset was an initiative relying on the participation of stakeholders:

> All countries and all stakeholders, acting in collaborative partnership, will implement this plan. We are resolved to free the human race from the tyranny of poverty and want and to heal and secure our planet. We are determined to take the bold and transformative steps which are urgently needed to shift the world on to a sustainable and resilient path. As we embark on this collective journey, we pledge that no one will be left behind. (UN General Assembly, 2015, p. 1)

The journey set up in the SDGs is a collective one—making it everyone's stake to progress and ultimately realise the set of ambitious goals. Furthermore, again, as it was the case in *The Future We Want*, this new partnership paradigm is rooted in the solidarity with the poorest:

> The scale and ambition of the new Agenda requires a revitalized Global Partnership to ensure its implementation. We fully commit to this. This Partnership will work in a spirit of global solidarity, in particular solidarity with the poorest and with people in vulnerable situations. It will facilitate an intensive global engagement in support of implementation of all the Goals and targets, bringing together Governments, the private sector, civil society, the United Nations system and other actors and mobilizing all available resources. (UN General Assembly, 2015, p. 10)

Here again, the document positions the SDGs as a monitoring programme produced with developing countries as key partners. Furthermore, the document posits the partnership as being one of a wider spectrum of such collaborations, involving national actors, the private sector and civil society. Thus, the SDGs become a participatory monitoring tool, requiring 'buy-in' in the broadest sense *and* pointing to consensus-building as the key underpinning principle of the new framework. Taken together, these two documents clearly show the underpinning logic of the SDGs as one of participation across institutional boundaries but also—and perhaps more importantly—across previously traditional lines of power and influence.

Of course, this does not mean that the introduction of this 'participatory' approach to statistics was straightforward—quite the opposite. Production of statistics is, in the end, not only a process aimed at consensus, but arguably predominantly a technical process following a specific set of methodologies. One way in which these democratic and technocratic ideals were married was in the concept of 'country ownership'. As a concept, it did not yield all the decision-making power to countries, but rather it attempted (not always successfully) to integrate political buy-in into the production of methodologically robust indicators. The technical group responsible for the indicator development—the IAEG-SDGs—has set up country ownership as one of their key goals. This positions this highly technical body as a broker of connections, rather than merely a methodological ombudsman:

> The role of the IAEG-SDGs members should include consultation and coordination within their own national statistical system, and should also include reaching out to the countries in their respective region and sub-regions. (IAEG-SDGs, 2015a, p. 2)

This point is further repeated in the discussion, as reported:

> During the discussion under this agenda item members of the IAEG-SDGs commented on the relationship between national, regional and global indicators, the need to ensure national ownership of the global indicator framework, the importance of statistical frameworks. (IAEG-SDGs, 2015a, p. 10)

The choice of focus on 'ownership' in relation to securing meaningful country participation is interesting here: on the one hand, it is malleable enough to bring together both the technocratic and the democratic logics of the SDGs. On the other, focusing on 'ownership' does not surrender the technocratic standards of the indicator development, but still communicates the need for countries to adopt the measurement and policy prioritisation of the indicators as political projects.

3 Democratic and Technocratic Logics in Action

The political declarations outlined in the key SDG documents and structures materialised in the ways the relationship between countries and IOs was designed and put in place. The principles of participation and technocracy—even though contradictory—were predominantly discussed as *indivisible*. Even though at the level of political declarations, some level of discrepancy was to be expected, the translation of these principles into specific measurement processes led to tensions and contradictions, particularly in various practices occurring at the intersection of work of experts and national policymakers and civil servants.

In particular, three settings were particularly prone to this dual logic of technocracy and democracy: (1) practices of securing country 'buy-in' into the monitoring frameworks; (2) practices of production of indicators on the country level requiring navigation between the overall standards of reporting and the local politics and finally, (3) practices of producing and using 'imperfect numbers'. Across these three types of settings—explained in detail in the remaining part of this chapter, the problem of navigating democratic and political accountability became not only visible but also actionable (Fontdevila & Grek, 2020; Grek, 2020). It required maintaining sufficient levels of technocratic accountability to reap the benefits of the 'epistemic virtue' (Daston & Galiston, 2007) of numbers (such as standardisation, objectivity and universality), whilst combined such virtue with important political calculations, such as securing consensus-building, collective action *and* the political acceptability of both the entire SDG framework and the specific numbers produced in the process.

3.1 Country Buy-In

One of the difficulties was the process of securing and maintaining country buy-in into both the monitoring framework of the SDGs but also to the specific indicators. The interviewed experts across the organisations saw it as a crucial process—at

times even more important than the more technical process of developing indicators themselves. This process (not unlike other quantified practices, see Chap. 3 on harmonisation) was happening in a two-way manner on the global and country levels.

On the global level, the key objective was to secure buy-in into the SDG framework itself. Here, the process of 'buy-in' was essentially a process of consensus-building. This required (at least for some indicators) a negotiation across the countries and IOs—and often even across the IOs themselves. An example here could be the negotiation of one of the most unique indicators within the SDG framework: 1.2.2 indicator of multidimensional poverty. This indicator was met with opposition from two directions: the countries as well as the World Bank.

The opposition of the World Bank was made on technocratic grounds. The key argument was one against measuring this target by using an index, aggregating different dimensions of poverty into one number. For example, as indicated during the second meeting of the IAEG-SDG:

> The selection of appropriate indicators for global monitoring depends on the interpretation of SDG target 1.2. If reducing poverty in every dimension is the major concern, then a dashboard approach—measuring each dimension of poverty/deprivation separately—would be an appropriate way to monitor progress at the global level. However, this would add a significant number of additional indicators to the framework, and since the SDG framework as a whole is a dashboard approach, introducing a smaller dashboard for SDG 1.2 could be confusing. Moreover, if the interest is to monitor the change in all dimensions of poverty using a single statistic, then there is an argument for considering a composite indicator—such as the MPI and others, which can be disaggregated to obtain the proportion of men, women, and children in poverty as required by the target. (IAEG-SDGs, 2015b, p. 3)

Multiple interviewees in the World Bank likened the Multidimensional Poverty Index (MPI) to be looking at a car dashboard: instead of tracking the level of oil or fuel in the car, the MPI was trying to 'summarize' how the whole car works using a single figure. This argument was made from a methodological standpoint and as such was based primarily on technical considerations.

The second source of opposition came from the countries that were rejecting what they saw as a limiting unitary approach to the measurement of multidimensional poverty. In particular, there was opposition to introducing the Global MPI as the countries thought it was not capturing their country-defined concept of poverty. The stakes were high, as the lack of 'buy-in' into the measure was seen as risking its complete removal from the framework. Eventually, the process was successful, thanks to negotiations with the relevant countries and compromises. As recalled by a United Nations Development Programme (UNDP) expert:

> A lot of it was about advocating for the measure in the SDG framework and that's an indicator that had some political pushback. So, in the initial meetings of the Inter-agency and Expert Group on SDG Indicators, our position was to advocate for it and to convey to the Member States the sense that if it was adopted as an indicator there would be significant support behind it, organisational support standing behind it and we did have partners with us that have been supporting that indicator, notably Oxford University, so OPHI in Oxford with Sabina, but as well as UNICEF, which has done quite a bit in also supporting countries in measuring multidimensional poverty. UNDP of course has developed a measure, that was way before my time, but UNDP has developed a measure and has been supporting the measure in quite a few countries and, to me, it's a measure that's important because it links quite closely with the policy priorities of a country, because it's just not a metric, but it's a metric that's about those policy issues that matter for a country and for the wellbeing of citizens in a country. (UNDP, 2)

Although the global MPI measure was eventually not introduced, the measure of the multidimensional poverty that was in the end included was one based on country-defined measures *without* any custodian IO agency taking responsibility for the measure; instead, countries themselves were to act as custodians of the MPI. Thus, the MPI represents a clear case of the level of political and technical compromising that the international community was prepared to succumb to, since this is the only target within the SDG framework that does not have an IO responsible for its implementation. This solution (advocated e.g., by Mexico) was seen as a consensus allowing countries to employ their own definitions and approaches to multidimensional poverty without the top-down methodological guidelines. This consensus—even though it secured the buy-in—was not seen as optimal from the perspective of the technocratic logic, as it risked measures that are not robust, comparable or indeed methodologically correct. As highlighted by an academic consultant:

> There are I can't remember how many, 369 targets in the SDGs? 368 of those targets have an international organization whose job it is to help national governments produce those measures and statistics. The 1 target out of the 369, where no international organization have any responsibilities to multi-dimensional poverty measure, I don't know why that is, but UNICEF does a bit of help, the World Bank does a bit of help. It's no one's job. Of course, you then get World Bank, UNDP, UNICEF, proposing different things, which is less than helpful if you're a national statistical office in Tuvalu or somewhere. Of course, the regional organizations try and also suggest different things, and academics suggest different things. We do hope he does. You end up with a nice selection, and no one [source]. […] There's a lack of technical help and expertise from international and UN organizations to governments to do something which they've never done before. Most countries have no experience in producing multi-dimensional poverty measures. They have, obviously, a statistical office that

> has professional ideas on what the quality should be of a national statistic. When they look at a lot of the proposed measures, they don't conform to their professional standards. That's a problem, so they'll resist. (Academic, 1)

On the country level, the process of securing buy-in into the indicators was equally complex. The twofold logic of the SDGs—drawing on both technocratic and democratic accountabilities—had to be navigated in the process of producing the measurement for specific SDG targets (Fontdevila & Grek, 2020). The interviewed experts working at country level were unified in their perception that the production of an indicator is at once a political, but also a complex methodological process:

> I think that the global conversation is a very different one from the country level conversation. The country-level conversation, nobody would disagree at the Bank if you were to say that building up poverty measures is a political process as much as a technical one. And then, and you see if you go into the details on how poverty lines are constructed across the world in which the Bank provided support, there's lots of variability in those. Whether some items are included or not included. How things are, whether prices are deflated across the country or not. All of those are like because in that particular country that country wanted this versus the other. (World Bank, 4)

Interestingly, even though the interviewees were not disillusioned about the politics of the measurement process around the MPI, they—with surprising uniformity—saw the central raison d'etre of the expert collaboration as providing 'technical assistance' to countries. As described by one of the interviewees:

> Typically, it's technical support, technical assistance that we provide to countries to compute MPI. We do that, UNDP does it, UNICEF as well does it quite a bit and it's technical support, it's capacity building and we've done some workshops, we've trained them on how to compute MPI and we've done also in collaboration with UNICEF some joint sessions to work together. At the country level, there is quite a number of countries where UNDP country offices work with governments and we're helping compute the MPI. (OPHI, 1)

Therefore, the key goal of the IOs working on the country level was to support the technical process of measurement and provide the capacity and knowledge base to fulfil the monitoring requirements. However, behind this 'basic' function was a much more subtle process of producing quantified knowledge that is politically acceptable (Bandola-Gill, 2021). Country 'ownership' therefore became a matter of navigating the tensions of usability of the indicators with achieving technical robustness: that is, produce indicators that reflect the reality but also are acceptable

to governments, fit the existing data structures, reflect trends and generally behave as all such global comparable numbers do:

> I think it boils down to how you can work with your counterparts to, in a way, get them to accept that this is what the evidence says, but also understand that they don't only have technical considerations, they have other considerations and work with them in terms of well how can this be useful to you. Maybe it's not the news you expected, but it's still the news, so what does this mean, you know, is there something that can be done proactively about this and so on. But it's not always easy and we have faced situations where the government didn't want to publish the numbers and the numbers have been not published or have been published with a delay. (World Bank, 5)

Here again, the strategy was to combine the technical advice (e.g., within the country report) with more specific political work, such as giving governments a 'heads up' about failing performance, exploring the potential risks, re-framing 'bad' numbers and generally performing functions that produce a politically and technically legitimate monitoring exercise.

3.2 Developing Indicators as a Participatory Process

The tension between technocratic and political accountability was further enacted on the country level when indicators were being developed and further adjusted. Similar to the global exercise, the process of developing country-level indicators was increasingly seen—by both experts and governments alike—as a participatory exercise. The indicators were produced through a collaborative process, which included the representative of IOs, government departments, academia and civil society or even members of specific populations affected by the process of measurement (e.g., the poor). The idea of 'technical assistance' provided by the IOs discussed in the preceding section when enacted on the country level went well beyond just scientific advice, and instead, it was a process of negotiation between multiple interests, ideas and objectives. As summarised by one of the interviewees, when it comes to developing the specific measures 'the process is more important than an outcome' (UNICEF).

This 'participatory turn' in the development of metrics was justified in multiple ways. On the one hand, the interviewees pointed to the issue of quality improvement of the indicator, as 'user involvement' might help to identify local issues or challenges which could help to make the indicator more robust. For example, as discussed by the following interviewee:

> I think that the challenge is it takes longer and you coordinate more people and they may have different views. But I think that the advantage is that if you think of who are the experts, well the experts are poor people who experience this […] I think it really sensitises you to that. So for example in El Salvador, talk about the challenges, the government made a trial measure with health. education, living standards and work. And then UNDP supported a two-year consultative process where people articulated their own deprivations. And El Salvador at that point was sort of the murder capital of the world, unfortunately. And the government said yes, but that's not poverty. But then there was an engagement and a dialogue and then by the end of that violence and esparcimiento [were included]. A place for like children to play, for the old people to drink coffee. (OPHI, 1)

Therefore, the interviewees saw these participatory processes as imbuing a technical process of developing an indicator—by its nature focusing on universalistic principles of science—with local meanings, ideas and politics. The predominant perception of this process was that it improved the measurement itself, as it allowed for a closer reflection of the reality of poverty.

On the other hand, interviewees mentioned more 'political' benefits of the participatory approach to developing the indicators. In particular, involving the wider spectrum of users was seen as increasing the political value of indicators through two means—by legitimising numbers and by improving their usability in political contexts. This idea was put forward for example by the following interviewee:

> With multidimensional poverty what matters more than anything is the process. And if you don't have the right partners at your table for conversation it's not going to be useful. It's not going to be used. You can do an index, you can do it with only NSO, the Institute of Statistics, fine, you do it. You have a new number to report every year but it's not going to make any change. Unless you have the right people. (UNICEF, 6)

Involving the stakeholders was seen as a way of assuring the legitimacy of the indicator—it was not only technocratically legitimate but also reflected the broader consensus of a wide variety of actors agreeing over the measure (Bandola-Gill, 2021). As indicated in the following quote:

> The other kind of participation which is vital is different parts of government and academia. Because by the time the measure is launched you want the government to own it, to be willing to act on it, understand it. You want the field leaders, the key idea, thought leaders in the country to understand it, otherwise when the press release comes, they might be caught off guard and try to discredit it. So, there's often also consultations with these other actors to make sure, that their input is gained so that measures really builds off their wisdom and knowledge and then also so that they understand it and support it and see how it could be useful to them. (UNDP, 2)

Furthermore—and perhaps even more significantly—involvement of the stakeholders in the indicator development process was seen as a way of assuring the use of an indicator by the policymakers. The co-production approach to indicator building was seen by the vast majority of interviewees as the best strategy of making the indicators 'usable' in policy, for example, introducing new policy programmes aimed at improving the indicators. Involving government officials in the development of an indicator raised the stakes for the indicator and was seen as a powerful motivating force to account for it in the governing practices (e.g., introducing the programmes to improve the indicator).

Nevertheless, despite these benefits of participatory approaches to the indicators' development process, the production of an indicator was not seen as entirely a user-driven process but rather was constrained by technocratic considerations. During meetings, the main structuring force of the agenda was the design of the indicator, hence the limits of the inclusion of the debate were in fact outlined by the methodology. This necessarily led to some conflict over the measurement and a need to navigate the democratic ideals with the technocratic standards. For example, the interviewees mentioned the female genital mutilation (FGM) as an issue that many of the stakeholders wanted to add as dimension of poverty. However, adding FGM into the model was rejected on the methodological basis of the need for the indicator to improve over time and irreversible procedures would render these indicators to be not responsive (or at least in the nearer future). Therefore, even though it was politically (and democratically) important, it was rejected on the technocratic grounds. Nevertheless, the civil society actors who saw the development an indicator as an access point to government and an opportunity to shape the political agenda. As such, even in cases that technical considerations prevailed, the participatory approach to measurement played a role in de-objectifying measurement (cf. Desrosières, 2015). Here, the goal was to continue the debate and keep issues on the agenda through contested indicator design, rather than either exclude them altogether or alternatively 'naturalise' them through their validation; either option would render them invisible and thus politically and technically less useful.

3.3 The Power of Imperfect Numbers

The final site of navigation of political and technocratic accountability was the production of 'quality' numbers. Thus far, it is becoming clear that the production of numbers is not happening despite politics or against politics, but rather that the modes of quantification embedded in the SDGs are from their inception shaped

irrevocably by this 'dual logic' of technocracy and democracy. The preceding sections have shown that this inherent tension has shaped the practices and principles of producing numbers for the SDGs. In particular, this section explores how the growing focus on the political value of numbers (and consequently a move away from pure technocratic modes of accountability and legitimation) opens up new possibilities and roles for 'imperfect' numbers. We look at three types of such numbers: ambiguous numbers, placeholder numbers and provisional numbers and the different political roles these 'imperfect' numbers play.

Ambiguous Numbers

The first type of imperfect numbers is the ambiguous ones. The democratic criteria for the development of the SDGs required a nuanced process of consensus-building and walking the tightrope between the country politics and the methodological principles of measurement. Here we can return to indicator 1.2.2 (of multidimensional poverty). The final wording of the indicator was:

> The proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions.

This was undeniably an acceptable consensus by the countries and IOs. Nevertheless, the wording of the indicators (combined with the lack of a custodian agency) did not 'solve' any existing conflicts over how to best measure multidimensional poverty but rather offered a way of avoiding the conflict altogether. In particular, the wording of this indicator was so open that different actors went as far as interpreting the meaning of the indicator 1.2.2. within the SDG differently. For example, the interviewees within UNICEF interpreted this indicator as a child-centric measure:

> It's not just a disaggregation. However, not everybody reads it that way. So, some people say no, that doesn't mean that we have to measure child poverty specifically, it means we have to disaggregate child poverty. So, there are some disagreements on how to interpret that particular indicator. We, we meaning the people in UNICEF that work with data, we have no doubt that this is what it is, and this was the intention. […] But the colleagues that have participated in that very adamantly said no, this is what we meant, and this is why it's different. And there was a lobby, and there is a whole coalition where UNICEF participates with NGOs that lobbied very hard to make this happen. (UNICEF, 8)

Contrary to this perception, the interviewees within Oxford Poverty and Human Development Initiative (OPHI)/UNDP and the World Bank interpreted this indicator as disaggregation of household-level data. This 'interpretive flexibility' (Sahay

& Robey, 1996) of the indicator has even further increased the breadth of the indicator to encompass multiple different forms of measurement as acceptable within the SDG framework. Consequently, different countries could continue carrying out existing measurements and IOs could continue promoting their own approaches to measurement while at the same achieving the multi-stakeholder consensus over the SDG1. Collective action in this context was dependent on the ambiguity of numbers: the latter was acting as a unifier of a diverse field of practices.

Here, the political function of this 'imperfect' number was to enable multiple meanings (and consequently values, agendas, ideas, etc.) within one monitoring framework. As such, this 'strategic ambiguity' (Sillince et al., 2012) of the indicator enabled it to act as a boundary object in the original meaning of the term (Star, 2010): that is, it allowed for different interpretations and actions between different groups, without necessarily solving the conflict amongst them.

Placeholder Numbers

A slightly different approach to accommodating 'imperfect' numbers was to use a number that although was not gathering full support, it could still be used as a 'placeholder' number (the usual term applied to such imperfect numbers by the expert community). These placeholder numbers were used as temporary solutions until the consensus over another—and improved—measure could be formulated. Therefore, this meaning of numbers was grounded in the assumption of the changeability of metrics, rather than their stability. These placeholder numbers were important enablers of political action, as they did not halt the political process and allowed to move on with other items on the agenda. One example of such placeholder numbers was the GDP, as summarised by an expert sitting on the IAEG-SDGs:

> We've said that we could only put things in at particular times because otherwise there are like 2,000 people who would like to have 2,000 indicators more. But we have said that we have the GDP as a placeholder. So, if they would have this fantastic number that they're talking about we could probably just make a switch, or if that isn't allowed then we could test because it takes a long time to make this anyway, so we could test it and then we could put it in by 2025 or. Like I've said to very many people who want to do things, if you have really good research studies or good analysis, nothing stops you from using the really tiny indicator that you have and when you're talking about that you just add this analysis and say, talking about sustainability and tourism, then we think that blah-blah-blah and if this thing would happen then it would be fantastic. (National Statistician, 1)

Such reliance on placeholder numbers has important implications. On the one hand, placeholder numbers have a productive role, since the 'ever-perfectability' of

numbers (Rocha de Siqueira, 2017) is a field of constant negotiation. Using a placeholder does not stop the debate over how to improve the measurement, so it is inherently generative of new ideas, approaches, connections and bodies of expertise, without endangering the technical process; on the contrary, it appears as strengthening robustness. On the other hand, using placeholder numbers can, perhaps remarkably, de-politicise numbers. As we know from the literature on evidence-based policymaking (e.g., Weiss, 1979), the calls for better-quality or 'perfect' evidence are often mobilised by political actors as a delay tactic, allowing to delay decision-making and retain the status quo.

Provisional Numbers

The final category of imperfect numbers is provisional numbers. These types of numbers are produced 'ad-hoc' using methodological shortcuts or using approaches that are drawing on rather provisional and imperfect datasets in the first place. Provisional numbers are used to ignite political action through their argumentative power: interviewees appeared aware that the numbers are convincing, even when they are not completely methodologically robust.

The underpinning logic of the provisional number is one of the argumentative power of numbers. Here, the interviewees acknowledged the importance of having one 'killer number'—one that could shock or embarrass the policymakers:

> I think that to me, the most important thing is a clear 'fact'. So it's like the fact that children are twice as likely to live in poverty as adults. That's the thing, and you can represent that in a figure if you want. You can just use the words. You can put it any way you like. But that's the, I think distilling the essence of complexity down to a real, a simple truth that can make people see things I guess differently than they've seen them before, but fundamentally understand something. (UNICEF, 7)

Having such a convincing and easily travelling 'fact' was not an easy feat—and such 'facts' were needed more often—and quicker than they could be produced. Sometimes such a number could have been simply 'guesstimated' in the meetings as long as they were used in a politically savvy way:

> Other's interest is: get me a number, even if it's an imperfect number. Look, if you want to measure it for impact at the level of policy, I completely agree anything could work, even a number that you just make up in the middle of the conversation to impress people. That's OK. (UNICEF, 2)

Nevertheless, these were the rare and most obvious examples of such strategies. More often, the numerical 'facts' were produced using valid methodologies albeit

ones that involved large doses of uncertainty to produce the key 'numbers'—particularly global ones, allowing for cross-country comparison. For example, the World Bank interviewees pointed to the need to negotiate between producing the comparable numbers (such as their global poverty number) and navigating uncertainty in employed methodologies, such as nowcasting. Here, the provisionality of numbers was more complex, as it was often veiled by the complexity of the statistical approaches to produce them:

> Now, we do have nowcasted numbers and forecasted numbers, so nowcast means if the last surveys are from 2018, we nowcast to the present, the forecast is into the future, but we call it as such. We are very careful to label it as such and so on. So, I think, again, this is also the bar being raised of number, of what we produce and I think it is a good bar to be raised because we also need to, firstly, we need to be responsible. It is because of these pressures [to produce numbers quickly] that the World Bank now produces numbers that are more current than they ever have been, but at the same time if you press it too much then the numbers have no meaning and I think there is a tension there.
>
> But yeah, because if you look back and go back and see how often off we are about the numbers because we are doing the best we can, then there is reason to be sober about it and right now we have actually, the one thing that has improved in the World Bank is that we do a very good job in communicating the numbers and we have learnt a lot from the national accounts folk that they come up with the GDP number and then they constantly keep updating it, going back to the past and updating it. So previously we thought oh, poverty is a tinderbox, very political, we need to be very sure about the numbers and then go out, but now we are a lot more fluid, we say that, OK, it is 9.1% right now and previously the number was 9.2. We do a good job of communicating it and I think what it has done is that it has, while there may be frustration the numbers are dated, there is trust also. (World Bank, 10)

In this context, experts navigated the uncertainty of the provisional numbers by claiming that it allows for greater transparency of the measurement process; second, they created opportunities for action, instead of the paralysing effects of trying to secure accuracy. To counter criticism, interviewees pointed to extensive methodological appendixes with clearly stated limitations of their studies. As such, the political argument for the role of provisional numbers was framed in terms of transparency and usability, and thus further strengthened the measurement process rather than detracting from it.

4 Conclusion

The epistemic infrastructure of the SDGs requires not only an extensive basis of data and indicators and their multiple inscriptions but also linkages between different actors, connecting different parts of the infrastructure. This chapter has explored the expansion of the connections and interdependencies grounded in the new governing paradigm of the SDGS—one of participation of all countries in the decision-making and design of this framework. As we have shown, this participatory logic was embedded in the SDGs from their inception with important consequences for the governance structures of this framework (see also Chap. 2) as well as the implementation of the framework on the country level. Consequently, the SDGs are driven by both the technocratic logic of quantification and the democratic logic of participatory governance. This chapter has illustrated how the newly emergent dual logic of the SDGs has resulted in new connections and linkages as well as creation of new political spaces of governing by numbers. Consequently, this turn to participation played an important political role in communicating the equality as the underpinning value of the SDGs as well as—or perhaps more importantly—securing the buy-in into the epistemic infrastructure of this measurement programme.

One of the underpinning assumptions of the epistemic power of quantification and its influence is the separation of the spheres of science and politics (Lahn & Sundqvist, 2017). And yet—this positioning of measurement as objective and devoid of politics is increasingly challenged not only on the grounds of ethics but also on the grounds of democracy, effectiveness and efficiency. It is increasingly acknowledged that quantification should strive not only for producing 'global' knowledge but also for acknowledging different contexts in which measurement is being done. These broader trends have been enacted and transformed within the Sustainable Development Goals. From their inception, the SDGs were designed to be both a highly technocratic monitoring programme of 'governing by numbers' (Miller, 2001) and a participatory project aimed at assuring the participation of countries and communities. As we argued in this chapter, this double logic of the SDGs permeated the practices of producing and using data on all levels of governance.

The focus on both democracy and technocracy has proved to be challenging for the experts within the IOs, as neither of the two logics could have been completely satisfied. Instead, the experts engaged in the process of 'sufficing' and navigating both logics and types of accountability: the technical and the democratic ones. This balancing act proved to be difficult, as prioritising either one of the two 'logics'

risks the loss of momentum and support: for example, as indicated in our discussion of different approaches to dealing with imperfect numbers, prioritising methodological practices of mechanistic objectivity (Daston & Galison, 2007) risked stalling collaborative action, politicising it or stopping the political processes aimed at actually fulfilling the targets of the SDGs. Alternatively, the technocratic process was mobilised when delaying practices aimed at changing the focus from often difficult political decisions, turned to seemingly endless and irreconcilable methodological debates. On the other hand, the baseline legitimacy of the process still rested on the epistemic virtues of numbers (cf. Bandola-Gill, 2021). Focusing entirely on democratic accountability risked inviting 'stealth' politics whereby the powerful actors got more influence within the 'participatory' processes.

This tension between democratic and technocratic modes of production of numbers (and their accountability) shaped the nature of quantification itself. Numbers are no longer 'fixed points' (cf. Lahn & Sundqvist, 2017) but rather more fluid entities that could be improved, changed and mobilised in different ways. Consequently, the underpinning logic of quantification changed—it is no longer a process of maximising the quality of numbers (in order to maximise the quality of political processes) but rather it is a process of multivocality (Bandola-Gill et al., 2021) where the quality of the political process is established by the quality of deliberation rather than facts that underpin it.

References

Bandola-Gill, J. (2021). The legitimacy of experts in policy: Navigating technocratic and political accountability in the case of global poverty governance. *Evidence & Policy, 17*(4), 615–633.

Bandola-Gill, J., Grek, S., & Ronzani, M. (2021). Beyond winners and losers: Ranking visualizations as alignment devices in global public policy. *Research in the Sociology of Organizations, 74*, 27–52.

Bandola-Gill, J., Arthur, M., & Ivor Leng, R. (2022). What is co-production? Conceptualising and understanding co-production of knowledge and policy across different theoretical perspectives. *Evidence & Policy*, 1–24 (online first) https://doi.org/10.1332/174426421X16420955772641

Best, J. (2014). *Governing failure: Provisional expertise and the transformation of global development finance*. Cambridge University Press.

Daston, L., & Galison, P. (2007). *Objectivity*. Zone Books.

Desrosières, A. (2015). Retroaction: How indicators feed back onto quantified actors. In R. Rottenburg, S. Merry, S. Park, & J. Mugler (Eds.), *The world of indicators: The making of governmental knowledge through quantification* (pp. 329–353). Cambridge

Studies in Law and Society. Cambridge University Press. https://doi.org/10.1017/CBO9781316091265.013

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). Routledge.

Fukuda-Parr, S. (2017). *Millennium development goals: Ideas, interests and influence*. Taylor & Francis.

Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures, 25*(7), 739–755.

Grek, S. (2020). Prophets, saviours and saints: Symbolic governance and the rise of a transnational metrological field. *International Review of Education, 66*(2), 139–166.

Hansen, H. K., & Porter, T. (2012). What do numbers do in transnational governance? *International Political Sociology, 6*(4), 409–426.

Inter-Agency and Expert Group on the Sustainable Development Goal Indicators (IAEG-SDGs). (2015a). *First meeting of the inter-agency and expert group on the sustainable development goal indicators New York, 1–2 June 2015. ESA/ST/AC.300/L3*. United Nations.

Inter-Agency and Expert Group on the Sustainable Development Goal Indicators (IAEG-SDGs). (2015b). *Open consultation for members and observers (As of 15 September 2015)*. United Nations.

Jasanoff, S. (Ed.). (2004). *States of knowledge: The co-production of science and the social order*. Routledge.

Jasanoff, S. (2011). The practices of objectivity in regulatory science. In C. Camic, N. Gross, & M. Lamont (Eds.), *Social knowledge in the making* (pp. 307–337). Chicago University Press.

Krick, E. (2018). The epistemic quality of expertise: Contextualized criteria for the multi-source, negotiated policy advice of stakeholder fora. *Critical Policy Studies, 12*(2), 209–226.

Lahn, B., & Sundqvist, G. (2017). Science as a "fixed point"? Quantification and boundary objects in international climate politics. *Environmental Science & Policy, 67*, 8–15.

Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. University of Chicago Press.

Miller, P. (2001). Governing by numbers: Why calculative practices matter. *Social Research, 68*(2), 379–396.

Nowotny, H., Scott, P. B., & Gibbons, M. T. (2001). *Re-thinking science: Knowledge and the public in an age of uncertainty*. John Wiley & Sons.

Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Rocha de Siqueira, I. (2017). Development by trial and error: The authority of good enough numbers. *International Political Sociology, 11*(2), 166–184.

Sahay, S., & Robey, D. (1996). Organizational context, social interpretation, and the implementation and consequences of geographic information systems. *Accounting, Management and Information Technologies, 6*(4), 255–282.

Sillince, J., Jarzabkowski, P., & Shaw, D. (2012). Shaping strategic action through the rhetorical construction and exploitation of ambiguity. *Organization Science, 23*(3), 630–650.

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values, 35*(5), 601–617.

UN General Assembly. (2012). Resolution adopted by the General Assembly on 27 July 2012. The Future We Want. A/RES/66/288.

UN General Assembly. (2015). Resolution adopted by the General Assembly on 25 September 2015. Transforming our world: the 2030 Agenda for Sustainable Development. A/RES/70/1.

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review, 39*(5), 426–431.

SDGs and the Rise of an Epistemic Infrastructure: Actors' Networks, Partnerships and Conflicts in the Education SDG

1 Introduction

The starting point of this book was an observation that despite the burgeoning number of publications on the global phenomenon of 'governing by numbers', our understanding of the relationship between the politics of measurement and how transnational governance comes into being is less well examined. Due to the fluidity and complexity of intense cross-boundary networks and 'soft' regulation regimes[1] that dominate the transnational space, transnational governance is a particularly productive space to interrogate the role of quantification as an infrastructure that facilitates governance (Djelic & Sahlin-Andersson, 2006). This chapter explores the role of different organisations and individual actors within them in creating the connections and outlining the directions of the new sustainability agenda.

By exploring the role of networks, meetings and interdependencies, we focus on the key process through which numbers, data and indicators are transformed into an epistemic infrastructure. The process of network-making that is at the heart of this chapter is as inherently a process of *infrastructuring* of the SDGs and creating common governing spaces around numbers. Thus, a central focus of this chapter is the examination of the politics of quantification through the prism of the political work of the actors (Lagroye, 1997) involved in the collection and validation of

[1] We are alluding to *soft law* here, which refers to officially ratified but not legally binding instruments like international entities' resolutions and declarations (of intent) etc.

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_6

country data as part of global performance monitoring (Grek, 2020b). By 'political work' we mean actors' practices undertaken in order to develop and manage the interdependencies between internal (upstream) and external (downstream) relationships, as well as personal ideas, values and interests;

> political work takes place through and across a range of configurations of actors who compete to construct alliances—*political enterprises*—that are capable of winning the negotiations they are involved in. (Jullien & Smith, 2008, p. 16; italics in original)

This chapter returns to the construction of the fourth of the Sustainable Development Goals (SDG 4) within the United Nations' 2030 Agenda (UNESCO, 2016). Having examined the discursive construction of the governing and measurement infrastructure of the SDG4 in "Scripting the SDGs: The Role of Narratives in Governing by Goals" chapter, this chapter will focus on the political work of actors involved in effecting and monitoring progress towards achieving SDG 4 by forming networks and collaborations around measurement. This work is complex and often contradictory, as it requires that they come together and share knowledge and working practices, whilst at the same time they need to preserve their respective organisation's unique 'brand', their expertise and services, as well as managing their own personal values and career aspirations. These centrifugal forces of collaboration and competition allow for an in-depth analysis of the workings of 'soft' regulation through quantification. Thus, examining these actors' work represents a unique opportunity to open a 'black box' in the field of global monitoring, rather than stacking yet another one (Bhuta, 2012).

2 Networks and Interdependencies in the Epistemic Infrastructures of the SDGs

The SDGs entail not only the extensive architecture of quantified measures but also—the system of connections and linkages that bring together the goals, targets, indicators and data into a complex web of relations, processes and practices (Tichenor et al., 2022). These linkages and interdependencies encapsulate a growing number of actors—from the UN agencies, member states to philanthropic and civil society organisations and academia. These complex relationships happen through and around numbers—which can both stabilise the connections and mobilise new constituencies and new interdependencies.

This chapter explores the SDGs as a particular form of network governance (Provan & Kenis, 2008) as they organise diverse social, political and economic

priorities around the process and practices of quantification. On the one hand, networks have been characterised as a key element of governance itself, since it is the collaborative arrangements between autonomous organisations, as well as the brokering between them, that make governance happen. Seen in this way, network coordination amongst diverse actors can enhance learning, use resources efficiently and deal with complex problems that require a multitude of perspectives and action (O'Toole, 1997). On the other hand, a focus on producing quantified knowledge as a key organising principle of these communities resembles what Peter Haas (1992) famously termed epistemic communities. This framework assumed that knowledge is being produced and shared in (international) policy through networks of experts. These experts have 'authoritative claims' (Haas, 1992, p. 3) to knowledge—entailing not only academic knowledge but also persuasive power over the interpretation of that knowledge. As argued by Haas (1992), the members of epistemic communities do not need to have the same disciplinary background, but they do have to share common values and beliefs, as well as the general understanding of causes and effects in specific areas and shared practices of validating knowledge. Unlike Haas' focus, the exploration of networks within the SDG governance points to broader diversity of participating actors, going beyond just elite experts and instead including different groups of practitioners, both from the Global South and North.

Empirically, this chapter traces the development of the production of the SDG4 in order to show the ways that the incremental build-up of the discourse, technical expertise and necessary—though always fragile—alliances facilitated a paradigmatic policy shift in the field of education: this is the move from the measurement of schooling (Barro & Lee, 1996) to the measurement of learning; that is, an emphasis on prioritising a focus on learning outcomes, skills and competencies, measured through what children 'can do' with the knowledge they acquire at school. In other words, instead of the traditional education statistics that measured inputs such as education expenditure, teacher salaries and length of the school year, the pendulum shifted to a greater interest in decontextualised, applied knowledge measured in real-life contexts. This chapter looks at the example of the SDG4 in order to illustrate the broader dynamics of the epistemic infrastructure. Although the work around the construction of the SDG4 (both prior to and after 2015) is not the only process that facilitated this shift (indeed its origins lie in New Public Management and the economisation of education discourse in the 1980s and early 1990s—see Gunter et al. (2016) and Ozga et al. (2009)), the global nature of the SDG4 process and the active involvement of most key education actors in its production led to a concerted effort to devise *global* learning metrics (Crouch & Montoya, 2019). Thus, the SDG4 became a prime metrological site of the production of this radical reconceptualisation of education measurement and policy with

implications globally (and not just across a handful of countries of the Global North).

3 From Schooling to Learning: The Emergence of an Epistemic Infrastructure

The discursive and metrological shift that moved the measurement agenda from a focus on schooling to learning began as early as 2000s. On the one hand, the OECD PISA, although measuring the skills and competencies of 15-year-olds in the global North (at least in the first rounds of the learning assessment and before its expansion in 2012 and 2015), received unprecedented media and policy attention worldwide; this was due to PISA's ranking of countries according to their education performance. PISA and subsequently the OECD prided itself in decontextualising education by focusing global, comparative testing not on the knowledge that students acquire at school (thus moving away from traditional ways of approaching schooling and curricula) but on what students can *do* with this knowledge. The OECD made direct links between countries' future competitiveness to how well schools prepare students to enter the labour market. PISA results were announced at the end of each testing cycle (every three years) and caused 'shock and awe' to many European countries in particular (and increasingly globally) including the 'education catastrophe' that hit Germany, or the 'education miracle' that turned Finland into an education tourist hotspot for education ministers and experts from around the world (Grek, 2009, 2013).

Nonetheless, perhaps more so than the OECD, it was the work of the World Bank that shifted the education debate, given the Bank's influence in the Global South (Prada-Uribe, 2012). The World Bank opposed the MDG emphasis on access to education, suggesting that lack of education had never been only a matter of whether children are in school or not; instead, it was suggested that the focus should be on what children achieve at schools when they are there. The work was undertaken by senior economists at the World Bank and the links to improved national economic growth were explicit from the start: in two key research reports (Glewwe, 2002; Hanushek & Kimko, 2000), it was suggested that individual mobility and better economic outcomes were achieved in countries that focused on knowledge and skills acquired in primary schools, rather than those systems that merely aimed to increase access. In 2006, another World Bank report became a milestone moment for education measurement, as it shifted the debate not only in education policy circles but also in development ones. The report, provocatively entitled 'From Schooling to Learning' (IEG-WB, 2006), was written by the

Independent Evaluation Group and created a polemical discourse against the MDGs' focus on access and completion: it suggested that the current emphasis was misplaced and that much more attention should be given to the improvement of skills and competencies, as it is the latter that leads to economic prosperity and better outcomes.

As a consequence, the Center for Global Development appointed three World Bank economists to further explore the issue; their report, *A Millennium Learning Goal: Measuring Real Progress in Education* (Filmer et al., 2006), unequivocally suggested that there was no evidence that showed that completion of primary school guaranteed the achievement of minimal levels of literacy and numeracy and that a re-think was long overdue. The materiality of data, reports and meetings intersected with the work of specific expert organisations and actors and led to a substantial policy shift, which was first taken up by specific governments.

Indeed, the arguments developed by the OECD and the World Bank had far more purchase in the development community groups, rather than in education (at least at the start). Both DFID (the UK's former Department for International Development) and USAID (the United States Agency for International Development) produced new strategies in the period 2010–2015 that identified the measurement of learning outcomes as an institutional priority and consequently channelled their education investments accordingly. Although there were a number of voices from academia that suggested that a singular focus on learning outcomes would take the attention away from other important pedagogical aspects (Barrett, 2011; Tikly, 2015), their commentary remained 'academic'; they had little policy influence and impact. Yet, there were still quite a few voices in education, especially those from UNESCO and the civil society, that were worried about the new trend and the misplacement, as they saw it, of education and schooling measures with those of outputs. Once again, the two functions of education, the humanistic and the economic one, were pitted against one another. The result was the slow emergence of 'a divide between those emphasizing quality and those primarily concerned with learning outcomes…Even if the differences between the two approaches were originally a matter of nuance or emphasis, they ended up forming two distinct communities of understanding, informed by different sets of ideas' (Fontdevila, 2021, p. 177)

As the decade progressed and the end of the MDG time frame was drawing to a close, we can observe a much more concerted effort not only to change the discourse (that had already been achieved) but also to start building an infrastructure for the establishment of a new measurement agenda, one in which learning, skills and competencies would be centre-stage and would replace the previous failing targets. The key protagonist in this new era was not the World Bank (though it was

always supporting in the background) but a new initiative, the Global Compact for Learning (GCL), which was launched in 2011 by the Brookings Institute Center for Universal Education. GCL quickly became an advocacy tool; through its reports, it created a sense of urgency, putting forward the idea that there was a learning crisis that was 'hitting the poorest, most marginalized and the youth particularly hard' (CUE, 2011). Just a year later, UNESCO in conjunction with the Global Education Monitoring Report (GEMR, 2012) published an estimate of the number of children not achieving basic literacy skills as reaching 250 million. The shocking figure became further ammunition not only for those that were pushing for the learning turn, but also for those who were suggesting the benefits of international learning assessments; without them, there would have been no evidence of this crisis. Thus, the crisis discourse had created a sense of urgency and would quickly turn into the need for *action*. Not only was it obvious that the MDG targets, set in 2000, were not going to be met, but also it had become evident—to some, at least—that these targets were ill-defined and misplaced and thus were failing millions of children around the world.

Crucially, GCL prepared the ground for the launch of another key initiative: the Learning Metrics Task Force (LMTF) was established in 2013 with the aim to 'catalyze a shift in the global conversation on education from a focus on access to access *plus* learning' (UIS/CUE, 2013; emphasis mine). This was a subtle, yet fundamental change and an open invitation to the two measurement camps to come together in search of the post-2015 agenda. Brookings invited the UNESCO Institute of Statistics (UIS) to head the task force, an important gesture towards an actor that appeared more trustworthy (to teacher organisations and civil society, at least) than the World Bank. More crucially, this was not an elite exercise; rather, LMTF was a very diverse organisation that included a wide range of actors not only from the International Organisations' expert world, but also from regional organisations, donors, governments, statistical agencies and civil society. The pluralistic nature of the membership, coupled with its UIS leadership and the timing (the preparations for the post-2015 agenda had already begun), made the LMTF the perfect opportunity to build the measurement infrastructure not only up but wide. This was the moment when the build-up of the new measurement agenda was to stretch across contexts and organisations to expand spatially, too. Essentially, the establishment of the LMTF became the foundation for building—what would later be called—the SDG4.

LMTF brought together a vast array of actors and organisations in its efforts to offer legitimacy to the task of shifting the debate and subsequently the post-2015 goals for education. As the previous section showed, it approached the contentious topic of the prioritisation of metrics and goals diplomatically, suggesting that they

were interested in exploring 'access *plus* learning' metrics. Thus, an olive branch was extended to academics, the civil society and professional organisations that perceived the learning focus as reductionist and as reflecting merely the economistic lens of the Bank's ideological positioning. Additionally, UIS' leadership (e.g., and not the World Bank's) gave the project not only credibility but also a ticket to move away from merely debating over priorities (the 250 million failing children was an alarm that kept on ringing) towards trying to find practical measurement solutions for their aims—in light of PISA and other regional, cross-national tests, the attention turned to the production of learning assessments, which, as it happened, have become the key data production machines for the SDG4 agenda (Fontdevila, 2021).

Despite the seemingly celebratory and ambitious language, the work of the LMTF was challenging, given that consensus had to be found not only on the aims themselves but also in relation to how these aims would translate into measurable indicators, as well as which spaces of deliberation would constitute the legitimate decision-making venues for making these choices. This is due to the fact that the efforts to devise the SDG4 indicator framework did not start by the UN Statistical Commission, but dated back to the establishment of an inter-agency, ad-hoc platform known as the Technical Advisory Group (TAG). Originally, the TAG was established by UNESCO in 2014 and recruited experts from UNESCO itself, but also from the GMR, the OECD, UNICEF and the World Bank. In many senses, while after 2014 LMTF 2.0—as the version came to be called—continued the debate at country level (Anderson, 2014), TAG adopted the work of the original LMTF with its focus on 'seven learning domains, and recommendations for global measurement areas' (Anderson, 2014). Chaired by the UNESCO Institute of Statistics, TAG was a much smaller grouping, with its membership limited to IO experts, and with the task to devise the 'post-2015' indicator agenda.

From March 2014 to May 2015, the TAG embarked on the process of mapping existing and potential education indicators, taking into consideration both their alignment with the (anticipated) targets and questions of data availability. Importantly, the work of the TAG benefitted from the input of a global consultation process, running from November 2014 to January 2015. In May 2015 the group's proposal was incorporated to the Framework for Action at the WEF in Incheon. That was a pivotal moment for the group's continuity, since the WEF recommended that the TAG is expanded, in order to include civil society and UNESCO member states organisations' representatives. It was partly the distrust towards the IOs leading the measurement agenda by the EFA actors, and partly the universalistic and participatory agenda of the SDGs that had brought this significant change, which also led to the renaming of TAG as the 'Extended TAG'. Subsequently, the Extended

TAG conducted ongoing open consultations led by regional leaders. Very quickly, what was a small, rather swift and efficient technical team of IO experts and representatives (with their own of course internal conflicts and competitions) had suddenly opened up to a much larger governing structure that required coordination, continuity, funding, support, meaning and a sense of purpose and unity: in other words, it became a complex infrastructure, ever expanding and changing, but always propping up and pushing the work of numbers.

Areas of concern for ETAG related to the issue of whether 'temporary placeholder' indicators should be devised, especially in relation to the lack of a universally comparable metric for learning outcomes. Above all, a major qualitative difference had already taken place in comparison to the previous education MDGs: five of the seven SDG4 targets now focused on learning outcomes and skills, a major departure from previous targets which focused on access and completion. In 2016, with the new SDG4 agenda formally adopted, the ETAG shifted again, giving rise to the Technical Cooperation Group (TCG), with the same broad membership (UIS, 2017) and remaining operative to date.

Additionally, in parallel to the TCG, another group came into existence, following in the footsteps of the LMTF: this was the 'Global Alliance for Monitoring Learning' (or GAML in short), the successor of the LMTF. Also created in 2016, GAML was originally defined as an 'umbrella initiative to monitor and track progress towards all learning-related Education 2030 targets' (UIS, 2016, p. 49), and was tasked with the development of tools, methodologies and shared standards to measure learning outcomes in the context of SDG4. Following the TCG, its membership is open to any individual or organisation willing to contribute to the work of GAML, and includes IOs, civil society organisations, a variety of technical partners and assessment organisations, and representatives of United Nations (UN) Member States.

Therefore, the political game of numbers became too high-stakes to leave it to the technical experts only. Even though the involvement of the majority of these actors, as the next section will show, was generally passive, the language of the new indicators became the new *episteme*—a way of knowing, describing and communicating about the world that was not merely about the craft of numbers (the *techne*) but involved the production of a new governing paradigm—a complex infrastructure of meetings, inscriptions, committees, agreements, supported by a growing number of networks of actors with different ideas, interests and alliances.

4 The Key Role of the Meeting

The transition from the ETAG to the TCG and from the LMTF to GAML were not without problems. Some original members of the TAG saw the TCG as a marker of the increasingly politicised nature of the indicators debate. At the same time, certain countries represented in the TCG and in GAML perceived that their input had not been sufficiently taken into consideration but simply used for rubber-stamping purposes. Others saw their role as primarily watchdogs, rather than full participants in the process. A civil society representative—involved in the TCG over a long period of time—elaborated on such tensions in the excerpt below:

> We were of course invited to be part of this, which was a clever move because we had probably been, if not the, at least one of the most critical voices in the room. So we had a dilemma and ended up actually agreeing to be part of this committee … I think what we struggle with is the fact that we know that just by being in the room, we are giving an indirect blessing of what the […] is doing. And at the same time, if we are not in the room, then we have no access to the conversations. We don't know what's going on. (Civil society, 1)

Thus, in this last empirical section, this chapter offers some observations on the procvess and practice of these groups' gatherings as the site where social, technological and material elements come together and stabilise these political spaces. Anthropologist Clifford Geertz's idea of the 'poetics of power' (Geertz, 1980) is useful for unravelling the thick layer of dramaturgy coating this apparently technocratic regime. Several of the study's interviewees suggested that most meetings are performative events, which follow a certain ritual, allowing enough free space to conclude with some loose decisions that determine the agenda for the follow-up meeting. The predominance of interviewees suggested that there is a clear-cut distinction between participants from the Global North, whose presence and contributions dominate the meetings, whilst representatives from countries of the Global South most of the time have a very passive presence, if any at all. This of course does not negate the agency and power of participants from the Global South, especially in relation to exploiting their own perceived weak positioning in order to accomplish specific goals.

Further, the ambiguity and informality of the process, despite being an issue for some in the room, becomes a valuable, malleable tool in ensuring participation, while at the same time also pushing on with a specific agenda. Interestingly, however, frustration and discord about the lack of transparency are not sufficient reasons to disassociate oneself from these alliances; being present at the discussions

even when one is at the receiving end is still considered more valuable than not participating in such meetings:

> Are we working on consensus basis? How do we deal with the fact that so many people have a conflict of interest? Who will draw conclusions? If there's voting, with what numbers would something have to be supported for it to be carried? And this was a frustration that grew as every session basically just ended with a broad sweeping, this was a very good discussion, thanks guys. And it was never really clear what anything would result in. (Civil society, 2)

Meetings are therefore key sites, where multiple levels of the infrastructure come together: material inscriptions in the form of data, PowerPoints, documents distributed prior and during the meeting, as well as the production and pursuit of common meaning and aims. A plethora of actors come together, irrespective of their own interests and ideas, in order to achieve a compromise, specific enough 'to keep the game going' (Bourdieu, 1990), but also flexible enough so that can be adapted and translated in their own contexts. Not everyone's participation has the same centrality and weight in these proceedings, nonetheless the expansion of actors is necessary in order to create some, even partial agreement and continuity (Grek, 2020a). As Luhmann suggested (1969), this is legitimacy achieved via procedure: meeting by meeting, compromise by compromise, the epistemic infrastructure achieves more than just the production of knowledge for policy: instead, it slowly shifts the needle towards a new, paradigmatic policy change. In the case of the SDG4, as we have seen, it was the fundamental policy shift from the measurement of access and enrolment data to the data and policy emphasis on learning outcomes. Chart 1 represents how some of these meetings and their actors are represented visually (orange squares represent meetings, whereas blue circles represent individual participants).

The visual depiction of how actors are connected through meetings is helpful here because it allows us to explore the ways that some actors are central (measured through their participation in most meetings) versus those that are more peripheral and those that might have attended only one or two meetings. Frequency of attendance denotes a more key positioning within the infrastructure, whereas less active participants are no less important; their inclusion and participation at least in some of the meetings adds legitimacy to the project and strengthens the infrastructure as a politically sanctioned operation. As clearly illustrated in this visualisation—the meetings played the key role in establishing the infrastructure by linking the actors, creating interdependencies between the organisations and establishing the stable flows of knowledge and power.

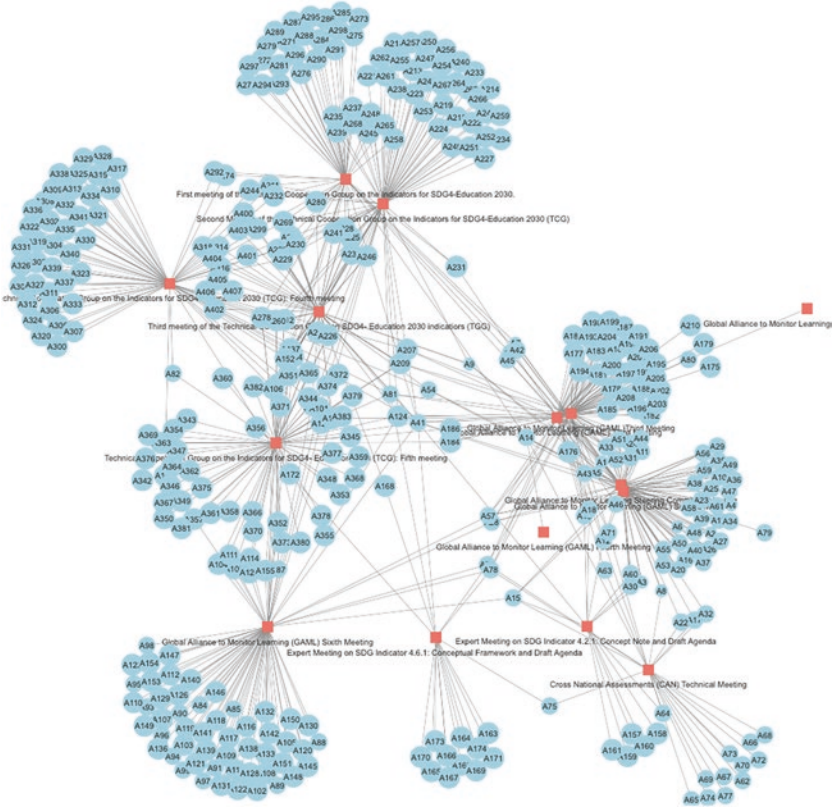

Chart 1 Social network analysis of SDG4 meetings

5 Discussion

This chapter has focused on the second order of the epistemic infrastructure (as outlined in the Introduction); that is, the role of interlinkages between actors in order to create not only a stable framework for measuring education but rather to establish it as a governing space in its own right. In order to make sense of the role of connections, networks and interdependencies in global public policy and their quantified enactments, one needs to highlight the main characteristics of SDG 4, as well as differences from previous similar initiatives. For a start, SDG 4 could be seen as a prime example of a transnational soft regulatory instrument (in the tradi-

tion of 'soft' law, i.e., best practices, expert standards, rankings, ratings, audits, quality assurance and the like). As such, it creates competitive and reputational pressures on those participating—by participants here we do not mean just country representatives, but also actors who work in the field, from the different IOs, research agencies and civil society. Similar to many other global monitoring tools, the Education 2030 Agenda (SDG 4) is interventionist in nature, leading effectively to a mutually constitutive relationship of the statistics and the countries they are meant to represent.

However, this chapter has also extensively discussed the ways in which SDG 4 is *substantially different* from other quantification exercises. The construction of SDG 4 represents a leap in the practice of transnational soft regulation in education because, although prescriptive, it also appears as transparent, pluralistic, open and developmental—consensus-making is prioritised and data collection and validation processes are required to be 'democratic'. "SDGs and the Politics of Reconciling the Dual Logic of Democracy and Technocracy" chapter discussed the ways in which the centrifugal forces of technical and political accountability have given shape to the epistemic infrastructure of the SDGs and the different actors' positionings and political work within it.

At the heart of this chapter are the paradoxes and the multiple ambivalences that quantification brings to the building of the epistemic infrastructure in education. On the one hand, they are necessary for the construction of discursive coalitions of actors who are not known to each other or have not collaborated before. Further, as we have seen above, SDG 4 identifies a specific failure of all previous statistical large-scale projects to deliver equitable education and develops a manifest governing programme to influence the behaviour of the participating actors. It may be that interventions still appear restricted to pushing (and largely financing) the statistical capacity for nations to produce data for governing; this step is seen as (and indeed is) key in achieving 'transformative' change. Thus, this chapter examines the promotion of SDG 4 as the emergence of an epistemic infrastructure that transcends the national/international/state/non-state divides. Crucially, this is not a governing structure that sits 'on top of' national and local decision-making, but has the very explicit aim of achieving national agendas capable of replacing current policies with ones which will deliver on the targets of SDG 4.

Nevertheless, these processes are not smooth and linear. As the empirical material shows, the process of establishing networks and interdependencies necessary for establishing the epistemic infrastructure involves antagonistic relationships of all the actors involved, and increasingly so, given the universal aspirations of the agenda and its claims to 'democratise' data monitoring for all the participant nations (as we have shown in "SDGs and the Politics of Reconciling the Dual Logic of Democracy and Technocracy" chapter). Lack of resources creates enormous

frustrations and limitations; in many ways it necessitates the use of pre-existing data. This creates pressures in the relationships of the four major IOs (UNESCO, OECD, UNICEF and the World Bank), since they have to coordinate their work in a context not only restricted by limited budget availability, but also under conditions of attacks on their expertise. How is one to make sense of how this complex infrastructure of transnational governance comes into being? In other words, in conditions where the policy area to be governed is fluid and constantly shifting, how is one to analytically explain how these multi-party, polycentric, transnational and often inter-cultural networks of governance function?

The ambiguity of numbers which describe and simultaneously prescribe allows participant actors to perform their function as transnational actors who can simultaneously take part in collective decision-making *and* maintain their own particular register of the meeting and its aims and decisions. Although the non-existence of any 'rules of the game' is often seen in the literature as an 'institutional void' (Hajer, 2003), where actors have to make up the rules and processes as they go along, this chapter suggests that quantification is precisely the necessary underlay in the construction of a relatively stable epistemic infrastructure where an emblematic issue, such as equal and quality education for all, brings actors together in an almost religious mission with strong moral undertones. Statistical knowledge in this instance, in the form of the specific targets and indicators constructed, despite being scarce, contested and sometimes altogether absent, still fulfils a key role in symbolically representing a much larger and complex political problem and serves to unite actors in the quest for 'solutions'.

References

Anderson, K. (2014). The learning metrics task force 2.0: Taking the global dialogues on measuring learning to the country level. Retrieved September 1, 2021, from https://neqmap.bangkok.unesco.org/wp-content/uploads/2019/06/The-Learning-Metrics-Task-Force-2.0.pdf

Barrett, A. M. (2011). A Millennium Learning Goal for education post-2015: A question of outcomes or processes. *Comparative Education, 47*(1), 119–133.

Barro, R. J., & Lee, J. W. (1996). International measures of schooling years and schooling quality. *The American Economic Review, 86*(2), 218–223.

Bhuta, N. (2012). Governmentalizing sovereignty: Indexes of state fragility and the calculability of political order. In K. Davis, B. Kingsbury, & S. Merry (Eds.), *Indicators as technologies of global governance* (pp. 132–164). Oxford University Press.

Bourdieu, P. (1990). *The logic of practice*. Stanford University Press.

Crouch, L., & Montoya, S. (2019). Measuring and reporting on SDG 4 and beyond: Lessons from the innovations in global learning metrics symposium. *UIS* (blog). http://uis.unesco.org/en/blog/measuring-and-reporting-sdg-4-and-beyond-lessons-innovations-global-learning-metrics-symposium

CUE. (2011). *A global compact on learning. Taking action on education in developing countries*. Center for Universal Education at Brookings.
Djelic, M.-L., & Sahlin-Andersson, K. (Eds.). (2006). *Transnational governance: Institutional dynamics of regulation*. Cambridge University Press.
Filmer, D., Hasan, A., & Pritchett, L. (2006). *A Millennium Learning Goal: Measuring real progress in education* (CGD Working Paper No. 97). Center for Global Development.
Fontdevila, C. (2021). *Global governance as promise-making. Negotiating and monitoring learning golas in the time of the SDGs*. PhD thesis. Facultat de Ciènces Polítiques I Sociologia, Universitat Autònoma de Barcelona.
Geertz, C. (1980). *Negara: The theatre state in nineteenth-century Bali*. Princeton University Press.
Glewwe, P. (2002). Schools and skills in developing countries: Education policies and socioeconomic outcomes. *Journal of Economic Literature, 40*(2), 436–482.
Global Education Monitoring Work. (2012). *Youth and skills: Putting education to work*. UNESCO.
Grek, S. (2009). Governing by numbers: The PISA effect in Europe. *Journal of Education Policy, 24*(1), 23–37. https://doi.org/10.1080/02680930802412669
Grek, S. (2013). Expert moves: International comparative testing and the rise of expertocracy. *Journal of Education Policy, 28*, 695–709. https://doi.org/10.1080/02680939.2012.758825
Grek, S. (2020a). Numbers as symbolic governance. *METRO*, 14 February [blogpost]. Retrieved June 3, 2020, from http://www.metro-project.eu/symbolic-governance/
Grek, S. (2020b). Interdependency in education governance. In G. Fan & T. Popkewitz (Eds.), *Handbook of education policy studies: Values, governance, globalisation, and methodology* (Vol. 1, pp. 309–328). Springer. https://doi.org/10.1007/978-981-13-8347-2_14
Gunter, H., Grimaldi, E., Hall, D., & Serpieri, R. (2016). *New public management and the reform of education: European lessons for policy and practice*. Routledge.
Haas, P. (1992). Introduction: Epistemic communities and international policy coordination. *International Organization, 46*(1), 1–3.
Hajer, M. (2003). Policy without polity? Policy analysis and the institutional void. *Policy Sciences, 36*(2), 175–195. https://doi.org/10.1023/A:1024834510939
Hanushek, E., & Kimko, D. (2000). Schooling, labor force quality and the growth of nations. *American Economic Review, 95*(5), 1184–1208.
IEG-WB. (2006). *From schooling access to learning outcomes. An unfinished agenda. An evaluation of World Bank support to primary education*. The World Bank.
Jullien, B., & Smith, A. (2008). Introduction: Industries, globalization and politics. In B. Jullien & A. Smith (Eds.), *Industries and globalisation: The political causality of difference* (pp. 1–28). Palgrave Macmillan.
Lagroye, J. (1997). *Sociologie politique* [Political sociology]. Presses de la Fondation nationale des sciences politiques (FNSP) and Dalloz.
Luhmann, N. (1969). *Legitimation durch Verfahren*. Suhrkamp.
O'Toole, L. J. (1997). Treating networks seriously: Practical and research-based agendas in public administration. *Public Administration Review, 57*, 45–52.
Ozga, J., Grek, S., & Lawn, M. (2009). The new production of governing knowledge: Education research in England. *Soziale Welt, 60*(4), 353–369.
Prada-Uribe, M. A. (2012). Development Through Data? A Case Study on the World Bank's Performance Indicators and Their Impact on Development in the Global South. A Case

Study on the World Bank's Performance Indicators and Their Impact on Development in the Global South (October 19, 2012). IRPA Research Paper, (5).

Provan, K., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory, 18*(2), 229–252. https://doi.org/10.1093/jopart/mum015

Tichenor, M., Merry, S. E., Grek, S., & Bandola-Gill, J. (2022). Global public policy in a quantified world: Sustainable development goals as epistemic infrastructures. *Policy and Society*. https://doi.org/10.1093/polsoc/puac015 (online first).

Tikly, L. (2015). What works, for whom, and in what circumstances? Towards a critical realist understanding of learning in international and comparative education. *International Journal of Educational Development, 40*, 237–249.

UIS. (2016). *First meeting of the Global Alliance to Monitor Learning (GAML)*. UIS. http://gaml.uis.unesco.org/wp-content/uploads/sites/2/2018/08/GAML-first-meeting-concept-note.pdf

UIS. (2017). Technical Cooperation Group on the SDG 4—Education 2030 indicators (TCG). Terms of Reference. http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2018/08/TCG_ToRdraft_20170929.pdf

UIS (UNESCO Institute for Statistics). (2016). *Laying the foundation to measure Sustainable Development Goal 4. Sustainable development data digest*. UNESCO Institute for Statistics. Retrieved June 9, 2020, from https://unesdoc.unesco.org/ark:/48223/pf0000245559

UIS/CUE. (2013). *Toward universal learning. What every child should learn*. UIS/Center for Universal Education at Brookings.

Expert Brokers: SDGs and the Emergence of New Forms of Expertise

1 Introduction

As carefully outlined thus far in the book, the space of global governance is constituted by various forms of knowledge practices and interdependencies between heterogeneous actors. Global governance is often described as a practice at the intersection of knowledge and policy (Haas, 2004). In this domain, these two spheres are particularly closely linked—where 'power is a disposition (in the sense of ordering or controlling) that depends on knowledge' (Adler & Bernstein, 2004, p. 294). The epistemic orders that dominate this space include technocratic modes of decision-making (Scheel & Ustek-Spilda, 2019) and quantification (Merry, 2016), which serve as ways of decontextualising knowledge spaces in order to create 'the global'—both a community that can work together and the common issues on which this community can focus. Nevertheless, global governance is never fully devoid of the 'local'—it is inherently a political space in which contextualised interests, actors and networks of power and influence overlap (Stone, 2019).

These spaces require 'readily comparable and accessible knowledge' (Rottenburg et al., 2015, p. 2). However, against the background of the highly fragmented structure of global governance a key question emerges: what forms and formats of knowledge fit these criteria? One approach is to focus on knowledge that has 'universalizing' qualities—such as numbers, standards and benchmarks (Timmermans & Epstein, 2010). These formats constitute global knowledge as they represent the 'view from nowhere' (Jasanoff, 2011)—knowledge that is context-less and as such is mobile and

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_7

applicable across different boundaries. Nonetheless, focusing on the 'view from nowhere' presents only part of the picture—particularly in the context of the growing paradigm of participation and democratisation of global decision-making, as reflected in the SDGs (see Chap. 5). Instead, global knowledge is increasingly characterised by what Mike Hulme (2010, p. 559) referred to as 'view from everywhere': 'knowledge which erases geographical and cultural difference and in which scale collapses to the global'.

Production of this type of knowledge is challenging as it spans multiple boundaries but also multiple knowledge orders. The 'machinery of knowledge production' (Knorr-Cetina, 1999) on the global level consequently encapsulates not the only production of research but—perhaps more importantly—unprecedented levels of coordination of actors, knowledge and practices. Within the SDGs, this machinery of knowledge production explicitly requires interdependencies between countries, International Organisations and supra-governmental groups—which we analyse as the second order of the epistemic infrastructure.

Against this backdrop, new forms of expertise emerge: as we have already seen, these include expertise in the production of narratives and visuals, the harmonisation of data and the promotion of participatory engagement and country buy-in. International Organisations are no longer just 'knowledge institutions' (Miller, 2007) producing numbers for global governance but are also boundary organisations (Guston, 2001), located at the intersection of different institutional, epistemic and political orders. As plentiful examples in this book have shown so far, the most important consequence of this evolution is the changing role of expertise: in this complex and fluid context, IOs are not only producers of numbers but rather *coordinators* of number production. They are 'expert brokers' whose role is not merely to produce knowledge but rather to create the conditions under which global knowledge can be produced.

Therefore, global governance as exemplified in the emergence of the Sustainable Development Goals is a unique space that is highly fragmented: hence, the processes of 'unification' (even if for short periods of time) of these diverse entities happen predominantly through knowledge practices. Consequently, the way of upholding the 'global' problems required in the SDGs is to create a common knowledge framework for knowing the problems. We have discussed this notion in Chap. 2; in this chapter, we turn to the actual role of actors who maintain these structures and practices. Numbers do not work on their own but rather they require a set of rhetorical and epistemic practices to make them operational. In this chapter, we turn to those practices undertaken by IOs to govern the SDGs and their key role in producing and maintaining the epistemic infrastructure.

The next section outlines the key debates in the literature on knowledge brokers and boundary work, which is followed by the analysis of the ways in which experts working in IOs become 'expert brokers'. We unpack this concept by focusing on three types of boundaries that these actors navigate: institutional, epistemic and praxis boundaries. We conclude this chapter by outlining the conditions under which global knowledge for sustainability is constructed.

2 Boundaries and Bridges

In the analysis of expert brokers, one of the central points of analysis is the boundaries between different knowledge systems. The concept of 'boundary work' was introduced by Thomas Gieryn (1983) as an approach to identifying the difference between science and other areas of human activity. Science, as argued by Gieryn (1983), is not identified by any essential, inherent characteristic but rather is demarcated by the rhetorical work of different actors as means of securing influence and resources. On the most pragmatic level, such a division helps to share labour between science and policy and to assign responsibilities for different elements of the science-into-policy process (Huitema & Turnhout, 2009). On the more conceptual level, such divisions play a role in differentiating between the 'technical' and the 'political', and therefore acting as lines of demarcation between knowledge and politics, fact and value, objectivity and interests (Turnhout et al., 2008).

Other scholarship has focused not necessarily on the demarcation but also on the navigation of the boundary (Halffman & Hoppe, 2004), hence assuming the flexibility and hybridity of the boundary (Epstein, 2011). Navigation of the boundary is one of the key practices necessary for the uptake of evidence in policymaking—which has led to a growing focus on the actors 'in-between' knowledge production and policy. The scholarship has explored different 'boundary bridges' (Wenger, 1998 see also: Kislov, 2014): knowledge brokers, boundary objects and boundary interactions. Knowledge brokers—individuals (or organisations)—are entities whose goal is to link different communities, translating knowledge between them and building capacity for cross-boundary engagement (Bandola-Gill & Lyall, 2017). They work on the periphery of two different social settings and are charged with enabling the collaboration and interaction between them (McNie, 2007; Miller, 2001). In order to achieve this, they must be perceived as grounded in and legitimate in more than one area of practice, as such they are accountable to both sides of the boundary (Guston, 2001). Finally, boundary interactions refer to different forms of connections between people from different domains, for example—meetings, networking, shared projects and collaborations. These

connections require diffusing the 'models of knowledge' (Lamont & Molnár, 2002) across different settings. Akin to Star and Griesemer (1989) view, the boundaries not only are markers of difference but also enable communication. Here, the focus is not on separating the practices but rather assuring their continuity across boundaries (Kislov, 2014). Building continuity across such boundaries is crucial to the functioning of the epistemic infrastructure as a whole.

Finally, these connections and continuities are often forged through indicators themselves, which serve as boundary objects and mediators in this global governance space. Star and Griesemer (1989, p. 393) describe boundary objects as those 'which are both plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites'. These objects have different definitions in different social worlds, yet they are recognisable across these disciplinary or governmental boundaries.

3 Expert Brokers: IOs and New Forms of Expertise

Perhaps the most significant finding coming out of the METRO project was the fact that, despite an undeniable commitment to governance by numbers, IOs did not see themselves as 'producers' of statistics. Rather, they saw themselves as 'links' between a variety of actors, navigating the complexities of knowledge production on the one hand and the politics of number-making on the other (see Bandola-Gill, 2021). This does not mean that the experts completely abandoned their statistical and economic training—quite the opposite, they perceived their work to be knowledge-intensive and requiring high-level skills. However, the central practice for experts working in IOs was not the production of knowledge but the *coordination* of knowledge production.

What we observe in this setting is an emergence of a specific form of expertise—one that surpasses producing and disseminating knowledge. Instead—this work requires creating specific knowledge environments in which knowledge is agreed upon and produced. This activity is not just translation (as discussed in the literature on knowledge brokers outlined in the preceding section): what these actors are doing is much more complex and involves creating the conditions for knowledge production through setting standards, coordinating their application and—finally—navigating the communication between these complex networks of actors. This process of coordination is complex, requiring bridging multiple boundaries at once—including boundaries between different institutional settings, epistemic boundaries between different forms of knowledge and differences in practices. These three types of boundary practices will be discussed in the remaining part of the chapter.

3.1 Brokering Knowledge

The first dimension of the process of creating global knowledge involved brokering knowledge. Even though experts working in IOs saw their role as predominantly supporting evidence-based policymaking—linking quantitative evidence with policy and practice—this process was more complex than translation between knowledge and action. Instead, the process involved navigating between different knowledge systems—from scientific, political and economic to local and indigenous and practice-based. This translation was multifaceted and required skills and integration of all different epistemic orders at once, which was challenging:

> This is a very country-specific question, it's just very specific to where are the levers, where's the interest in government, who are the partners? You need to be able to work within that realm to find your way to what is going to be most effective. So I think we just want to find a balance [with strict guidelines]. But it's really encouraging that balance, but within some parameters which fall within: 'yes this is child poverty measurement'. (UNICEF, 3)

On the country level, this translation between knowledge of different epistemic standing (c.f. Bandola-Gill, 2019), for example, between experiences of poor people, statistical and economic methodologies and political focus on 'doability', required embedding the numbers in other elements of the existing knowledge system—for example, translating them into storylines (see Chap. 4), linking them to existing agendas, policy solutions, etc. The key epistemic challenges the brokers must navigate are the different understandings of 'numbers' between experts and policymakers, including the methodological and technical standards and political context:

> A lot of times I feel, and this is something that we've talked about with our teams very frequently, that our role is not just to provide the best possible technical advice but it's also to work with our counterparts to think about how can they communicate that in simple understandable terms to the population or to other parts of the government who may not be experts in poverty measurement. So yeah, just being aware of what's the political context in which these changes are taking place is very critical. (World Bank, 12)

Here, opening the channels of communication was the key practice. As recalled by one of the UNICEF members—'the worst thing you can do is to just drop the report on their lap' (UNICEF, 7). The role of the brokers was to get the policymakers involved in setting the priorities, choosing the dimensions and shaping the

measurement process. Importantly, another key area of expert brokers' work was to keep the policymakers abreast of the numbers that are coming—particularly if the numbers were negative. As such, the co-production of quantified projects—in which the government, IOs, statistical offices and civil society representatives worked together to produce the measures and the reports together—became a space of boundary interaction in which these communities, previously fragmented, were becoming a community of practice.

On the global level, the process of consensus-making and coordination between different ways of knowing was complex. One specific site requiring navigation between different ways of knowing in order to achieve the 'view from everywhere' was a negotiation around the SDG reporting (Jasanoff, 2017). In the case of global poverty, this process was particularly complex, as the interviewees saw poverty as a situated and country-defined issue. This was an area in which different knowledge systems were competing. On the one hand, different countries as well as IOs were proposing poverty measurement more closely aligned with their understanding of poverty. Unsurprisingly, different organisations called for the inclusion of dimensions of poverty aligned with their organisational remit. These additional dimensions were not inconsequential as they opened up multiple possibilities for the custodian agency in charge of the reporting for this indicator. For example, UNICEF's written comments in the consultation on the SDG indicators submitted for the first meeting of IAEG-SDGs as *Indicator proposals received from agencies* (IAEG-SDGs, 2015) suggested the addition of child poverty but also suggested themselves alongside the UNDP as custodians of the target (based on the scale of coverage by both organisations):

> Proportion of children living in multidimensional poverty. This indicator is expressed as a percentage. Deprivation dimensions and indicators should be based on internationally agreed standards and definitions. Deprivation dimensions include inter alia: nutrition, education, health, housing, water and sanitation. (UNICEF)

The International Labour Organisation (ILO) requested employment status and saw themselves (alongside the WB) as the custodians. The International Fund for Agricultural Development (IFAD) called for an addition of disaggregation between rural and urban areas. The working group on disabilities wanted a disaggregation between able and disabled. The International Finance Corporation (IFC) asked for an addition of another indicator to this target:

> Percentage of population using banking services. Please disaggregate by gender. (IFC)

Correspondingly, the UN WOMEN requested women-centred indicator:

> Proportion of people who have an independent source of income by sex, age and source of income. (UN WOMEN)

These debates went beyond the content of indicators but also accounted for methodological standards. The best example here was the position of EUROSTAT:

> MPI should have the form of the EU 2020 poverty and exclusion indicator: if a person suffers from any dimension of poverty s/he has to be considered 'poor'. Statistical compensation of poverty dimensions against each other should not be allowed. Dimensions should be oriented towards basic needs: enough clean water, enough healthy food, clean air, shelter, security, basic education. (EUROSTAT)

On top of these debates over the understanding of poverty through disciplinary and organisational lenses, different countries were opposing measures that were not aligned with country-level definitions, as one interviewee put it:

> For multidimensional poverty, I think the problem is it's more acute, because countries are very different, so depending on how you choose the dimensions, some dimensions would be relevant in some context and not relevant in other contexts, so when you take dimensions or indicators that, for example, look at access to technology or access to a TV, which was the case in some of these indicators, you will find that in some societies everybody has that basically, it's a given, whereas in some societies there's still a significant gap in access to that kind of technology. So, if you use that in a setting where everybody basically has, let's say everybody has internet and broadband, you can't use that as a measure of poverty, because then nobody would be poor in that country. […] That's one of the reasons why you had this pushback by some Member States who didn't want to have a global definition that would determine the number of poor people in their country and that may not necessarily reflect the experience of poverty in their countries. (UNDP, 4)

It soon became clear that achieving the perfect consensus ('view from nowhere'—Jasanoff, 2017) was not achievable in this context, At the second meeting of IAEG-SDGs, the representative of Mexico's CONEVAL (The National Council for the Evaluation of Social Development Policy) advocated for a country-led indicator:

> Based on the foregoing, it is proposed that the indicator for Target 1.2 traced for Objective 1 could be stated as: 'By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions.' This proposal comprehends two essential aspects of Target 1.2: it i) recognizes the multidimensionality of poverty, and ii) favours each country with their specific needs. Also, allowing each country's particular definitions would take into account the specific contexts of poverty and the availability of sources of information in each one of them. (Mexico's country statement)

This proposal was challenged by Sabine Alkire—who works for OPHI and is one of the co-creators of the MPI—in her 'Academic Stakeholder Statement' in which she argued that only a global MPI would provide a comparable measure. She claimed that relying on country measures will not benefit the poor—because the national measures are not comparable and consequently in some countries, this approach will result in a very limited anti-poverty action.

In the end, as we have already seen in Chap. 5, the approach discussed by Mexico was indeed implemented and eventually, the country-led measurement was introduced to the framework. Therefore, in the end, the consensus in this case led to prioritising country-level data—or 'view from here' rather than the preferred option of global-level data. However, the alignment of actors was seen as a priority over finding 'perfect' data solutions.

3.2 Brokering Practice

The second boundary that was 'brokered' by the experts working in IOs was one of practices—including workshops, training sessions, online learning and handbooks aimed at capacity building. These strategies were not just aimed at creating specific knowledge repertoires for the stakeholders to use but they have a more important role in creating—what Galison (1997) called the 'rules of exchange' within this particular 'trading zone' (Collins et al., 2007). These rules of exchange allow for coordinating and collaborating between different actors without necessarily requiring the consensus.

For example, in the case of global poverty, this meant that economic and statistical methods were constructed as a standardised framework for exchange—of ideas, priorities and items on the agenda. This was not a mere capacity building exercise, as it was providing language through which the participants could communicate—the participants were expressing their ideas and priorities in the language provided by the methodological framework. As we discuss in Chap. 4, this common language provides a narrative for the global public policy of sustainability, providing a 'glue' for the epistemic infrastructure. Interestingly—and relatedly—the process of developing capabilities during the workshops was not completely open—the boundaries were established by the methodology itself, even though there was a level of flexibility:

> Stakeholders, I mean, you have to facilitate to them. Of course, because this [multidimensional poverty] is a new concept to many of them. In a way, we try to be very honest, we try to make the stakeholders to believe in us and try to make them reach a consensus that we want them to reach. Just to put it very, very frankly. Of course, we

> design all these workshops and materials, we give presentations, we convince them. This is firstly important. Secondly, the methodologies we are using are very relevant and they're technically sound methodologies. And then we try to engage them to identify what are the most important dimensions, but even for that, we need to come up with a list. As far as I recall, we did not let people brainstorm. We already had a list. Then people went through the list and say, okay, they select some. Although, I think in the case of Botswana, actually, they added something. […] As you can imagine for African countries, crime safety is a problem, so they want to add a crime safety dimension to it. And we have had no problem with that. If they think this is an important aspect of possible deprivation for children. So, we added that. (UNICEF, 10)

However, IOs' experts were aware that capacity building—with its designated roles for certain actors—was not necessarily aiding progress towards achieving the SDGs:

> Policy's really hard, because even if we think of a country with just a well-functioning data architecture and system in place, if I think of the US, it has excellent data sources, government technicians are very strong, collaboration between academics and government and policymakers are probably some of the best in the world. And yet, I'm not sure I could point to specific examples of data that's been collected that has directly influenced policy. (World Bank, 2)

In many cases, capacity building was not only a matter of knowledge creation but also a matter of getting all the actors to get the common ground and thus enabling communication. As we discussed in Chap. 3, statistical capacity development made explicit how central creating these conditions of knowledge production is to the functioning of the epistemic infrastructure of the SDGs. By mapping out and facilitating the materialities needed to expand, create and maintain data and statistical systems, IOs function as brokers for the entire system.

4 Brokering Global Public Policy: Boundary Work as Infrastructuring

Finally, we turn to the brokering role of numbers themselves. The predominant outlook in the literature on indicators (Merry, 2016) posits numbers as effective forms of both communication and governance as they create mobile knowledge which travels across different contexts. This aspect of numbers is indeed important and evidently, the SDG reporting framework draws heavily on this 'universality' of numbers, allowing for articulation of common goals within one monitoring scheme. Nevertheless, our findings point to another, more *situated role of quantification*—

namely the ability to create fora for debate and consultation amongst multiple stakeholders. This 'new' role of numbers in this context was made possible by the acceptance of quantification as a governing paradigm (Tichenor et al., 2022), as best illustrated by the fact that, despite ranging views of SDGs themselves, none of our interviewees questioned the idea of measurement itself. One of the interviewees summarised this paradigm (attributed to Andreas Schleicher, Director of Education and Skills at the OECD): 'measuring the pig doesn't make it fatter, but at least it can tell us if it's overweight or underweight'.

At the same time, the interviewees were unified in their perception that it was the process of producing the numbers—negotiation, consultation and consensus-building—which was as important as the numbers themselves. This unprecedented focus on the process of collaboratively producing numbers in the global governing spaces highlighted two interlinked qualities of numbers in this setting. First, numbers do not exist independently of the social actors producing them. In the context of the SDGs, this increasing coupling of actors and numbers led to the rising importance of the process of selection and coordination of multiple stakeholders. This makes the role of a broker not only one of a connector (Meyer, 2010) but also one of an architect—by shaping who gets invited to the process and who has a say (as will be illustrated in the example of meetings)—key to constructing the larger epistemic infrastructure of the SDGs. Second, numbers do not only represent the issues and communities but they also *create* them. As such, the process of quantification emerges as a process of 'stakeholdersation' (Metzger, 2013)—where indicators transform groups of actors into stakeholders of issue that are being measured. These two qualities of 'brokering connections' will be discussed in the remaining part of this section.

4.1 Numbers as Mediators

One of the key effects of the indicators, as identified by the vast majority of interviewees, was that they enable collective deliberation and decision-making. This ability to 'bring people into the room' (UNICEF 2, 4) was seen as central in this process as it opened up—often cross-sectoral—channels of communication and constituted a strategy for establishing common ground. This quality is central in creating cross-boundary initiatives (Bechky, 2003) and in the context of the SDGs it was achieved through quantification. This was particularly important for settings where the group structure was difficult to be maintained. One such area is poverty

which is a multidimensional phenomenon (Atkinson, 2019). This makes it a difficult policy concept, as the policymaking happens within organisational silos.

The development of multidimensional poverty measurement enabled the cross-sectoral engagement on a scale that surprised even the creators of the measures. As recalled by one of the interviewees, the most unexpected consequence of the development of the multidimensional poverty measurement was the fact that it brought various groups of stakeholders together and enabled deliberation. The interviewee argued that the indicator brought together departments which usually do not work together, even at the national level:

> Often ministers compete with each other: 'I want to be the most important minister, I want the most budget'. And when you have an MPI and the minister sit at the table then you can't move the MPI down the field single-handedly, you need a team. And so, you need to kick from the minister of health, kick from the education and together, as a team effort, they can move the ball on property. And they learned that. So actually, you might say 'I'm sending an education indicator to the education minister responsible for that'. But then she will say, 'I cannot make my education goal without the other ones'. So actually, they learn about what the others are talking about, how we need to integrate the policies, but they learn it from each other. And we've seen it in the number of governments where ministers actually reach a common understanding. Well maybe there they compete in other spheres but when it comes to poverty the moral imperative is so great that then they said we going to cooperate. (OPHI, 1)

Therefore, the indicators themselves have specific 'constituencies' (Voß & Simons, 2014)—sets of actors which emerge to maintain or develop specific tools as modes of governing. Even though this quality of the multidimensional poverty indicator was initially seen as an unintended consequence of the methodological innovation, it soon became common practice to mobilise it strategically. The central affordance in which indicators as material mediators shaped the practices was through their impact on practices. The practice of producing the indicator not only focused on the indicator itself but also created specific knowledge/knowing spaces:

> I think my observation here is we need to engage people at the right level, as high as possible, and do good facilitation. Don't just let them brainstorm. It's not going to be helpful because they don't even have the basics of what we are going to do. But do give them enough space to talk about what they think would be most important. […] Also actually, we also involved—I think we involved the academia from the country. So, we involved some university people to be with us in the facilitation, so that we are not really coming up with crazy things that are not relevant. But we make sure that these are relevant, and they're endorsed by local researchers, as well. I think that would be important. (UNICEF, 7)

Even though the process of developing the indicators supported the creation of constituencies, the process of generating global knowledge required more extensive coordination of different actors. The process of 'production of numbers' for global governance is in fact a complex navigation between these multiple actors who are producing the measures:

> I think it boils down to how you can work with your counterparts to, in a way, get them to accept that this is what the evidence says, but also understand that they don't only have technical considerations, they have other considerations and work with them in terms of well how can this be useful to you. Maybe it's not the news you expected, but it's still the news, so what does this mean, you know, is there something that can be done proactively about this and so on. But it's not always easy and we have faced situations where the government didn't want to publish the numbers and the numbers have been not published or have been published with a delay. (World Bank, 9)

On the global level, as we discussed in Chap. 2, these indicators mediate the frameworks and needs of other diverse groups of actors. As IOs create spaces for deliberating the international concept and the measurement tools for each indicator—as part of the required protocol for establishing or reclassifying an indicator for the framework set by the IAEG-SDGs and the UN Statistical Commission—civil society groups, member states, regional organisations and IOs provide both conceptual and methodological feedback to shape this unifying object of concern for the specified policy arena. As was the case with the indicator on governmental migration policies outlined in that chapter, almost 300 governments, International Organisations, civil society groups and academics were brought together by the process, allowing the indicator to become the common language spoken within such a vastly different group of stakeholders in shaping what is knowable about migration. Each indicator does this work of brokering diverse actors and perspectives in the infrastructuring of global public policy.

5 Conclusion

This chapter began with the acknowledgement of the complexity of global governance. It has become increasingly fragmented, and this fragmentation necessarily requires a form of linking, of collaboration and of creating common deliberative spaces. Consequently, governing these spaces requires creating specific conditions for knowledge production which become constructed not only by a specific set of indicators but also by the act of governing itself—the ways in which experts become

not only producers of knowledge but also—or even exclusively—knowledge governors. They do not produce numbers—they broker connections, build consensus and work towards the shared meanings, rules of exchange and material artefacts.

The scholarship on knowledge brokers sees them predominantly as actors 'in-between' whose role is supporting evidence translation between research and policy (Meyer, 2010; Turnhout et al., 2008). Even though this was one of the types of practices that the experts in IOs engaged in, their role was in fact more extensive. By 'bridging' different groups of actors, knowledge and practices, the expert brokers not only engaged in linking different epistemic and institutional orders but rather supported creation of unique forms of 'global knowledge'. The process of creating the conditions for knowledge production was not merely an act of translation but rather an act of creating new epistemic environments which then determined the forms of knowledge that were possible to produce. As such, these brokers acted as institutional filters, transforming a multiplicity of ways of knowing, organisational structures and political priorities into common epistemic and political frames. This was particularly important in a context of the high fragmentation of global governance—the role of brokering work was to create conditions for unifying of knowledge, without necessarily universalising it.

The key strategy here was to create specific 'fora' for engagement—for example, through meetings, indicator development or training workshops. The role of brokers was to bring different actors together to deliberate and build consensus. It was in these fora where 'the scale collapses to global', even if temporarily. The production of global knowledge was therefore a process of bringing actors together and letting them go back to their 'main' institutional settings and through this process allowing for the co-existence of multiple, sometimes contradictory, knowledge systems.

References

Adler, E., & Bernstein, S. (2004). 13 Knowledge in power: The epistemic construction of global governance. *Power in Global Governance, 98*, 294.

Atkinson, A. B. (2019). *Measuring poverty around the world*. Princeton University Press.

Bandola-Gill, J. (2019). Between relevance and excellence? Research impact agenda and the production of policy knowledge. *Science and Public Policy, 46*(6), 895–905.

Bandola-Gill, J. (2021). The legitimacy of experts in policy: Navigating technocratic and political accountability in the case of global poverty governance. *Evidence & Policy: A Journal of Research, Debate and Practice, 17*, 615–633.

Bandola-Gill, J., & Lyall, C. (2017). Knowledge brokers and policy advice in policy formulation. In M. Howlett & I. Mukherjee (Eds.), *Handbook of policy formulation* (pp. 249–265). Edward Elgar Publishing.

Bechky, B. A. (2003). Sharing meaning across occupational communities: The transformation of understanding on a production floor. *Organization Science, 13*(3), 312–330.

Cetina, K. K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.

Collins, H., Evans, R., & Gorman, M. (2007). Trading zones and interactional expertise. *Studies in History and Philosophy of Science Part A, 38*(4), 657–666.

Epstein, S. (2011). Misguided boundary work in studies of expertise: Time to return to the evidence. *Critical Policy Studies, 5*(3), 323–328. https://doi.org/10.1080/19460171.2011.606306

Galison, P. (1997). *Image and logic: A material culture of microphysics*. University of Chicago Press.

Gieryn, T. F. (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review, 48*(6), 781–795.

Guston, D. H. D. (2001). Boundary organizations in environmental policy and science: An introduction. *Science, Technology, & Human Values, 26*(4), 399–408. http://www.jstor.org/stable/10.2307/690161

Haas, P. (2004). When does power listen to truth? A constructivist approach to the policy process. *Journal of European Public Policy, 11*(4), 569–592. https://doi.org/10.1080/1350176042000248034

Halffman, W., & Hoppe, R. (2004). Science policy boundaries: A changing division of labour in Dutch expert policy advice. In S. Maasse & P. Weingart (Eds.), *Scientific expertise and political decision making* (pp. 135–152). Kluwer.

Huitema, D., & Turnhout, E. (2009). Working at the science–policy interface: A discursive analysis of boundary work at the Netherlands Environmental Assessment Agency. *Environmental Politics, 18*(4), 576–594. https://doi.org/10.1080/09644010903007427

Hulme, M. (2010). Problems with making and governing global kinds of knowledge. *Global Environmental Change, 20*(4), 558–564.

IAEG-SDGs. (2015). *Indicator proposals received from agencies* (May 2015). First meeting of the IAEG-SDGs, June 2015.

Jasanoff, S. (2011). The practices of objectivity in regulatory science. In C. Camic, N. Gross, & M. Lamont (Eds.), *Social knowledge in the making* (pp. 307–337). Chicago University Press.

Jasanoff, S. (2017). Virtual, visible, and actionable: Data assemblages and the sightlines of justice. *Big Data & Society, 4*(2), 2053951717724477.

Kislov, R. (2014). Boundary discontinuity in a constellation of interconnected practices. *Public Administration, 92*(2), 307–323. https://doi.org/10.1111/padm.12065

Lamont, M., & Molnár, V. (2002). The study of boundaries in the social sciences. *Annual Review of Sociology, 28*(1), 167–195.

McNie, E. C. (2007). Reconciling the supply of scientific information with user demands: An analysis of the problem and review of the literature. *Environmental Science and Policy, 10*(1), 17–38. https://doi.org/10.1016/j.envsci.2006.10.004

Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. University of Chicago Press.

Metzger, J. (2013). Placing the stakes: The enactment of territorial stakeholders in planning processes. *Environment and Planning A, 45*(4), 781–796.

Meyer, M. (2010). The rise of the knowledge broker. *Science Communication, 32*(1), 118–127. https://doi.org/10.1177/1075547009359797

Miller, C. (2001). Hybrid management: Boundary organizations, science policy, and environmental governance in the climate regime. *Science, Technology & Human Values, 26*(4), 478–500.

Miller, C. A. (2007). Democratization, international knowledge institutions, and global governance. *Governance, 20*(2), 325–357.

Rottenburg, R., Merry, S. E., & Park, S. (2015). In J., & Mugler, J. (Ed.), *The world of indicators: The making of governmental knowledge through quantification*. Cambridge University Press.

Scheel, S., & Ustek-Spilda, F. (2019). The politics of expertise and ignorance in the field of migration management. *Environment and Planning D: Society and Space, 37*(4), 663–681.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, "Translations" and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science, 19*(3), 387–420. https://doi.org/10.1177/030631289019003001

Stone, D. (2019). *Making global policy*. Cambridge University Press.

Tichenor, M., Merry, S. E., Grek, S., & Bandola-Gill, J. (2022). Global public policy in a quantified world: Sustainable development goals as epistemic infrastructures. *Policy and Society*. https://doi.org/10.1093/polsoc/puac015 (online first)

Timmermans, S., & Epstein, S. (2010). A world of standards but not a standard world: Toward a sociology of standards and standardization. *Annual Review of Sociology, 36*, 69–89.

Turnhout, E., Hisschemöller, M., & Eijsackers, H. (2008). Science in Wadden Sea policy: From accommodation to advocacy. *Environmental Science and Policy, 11*(3), 227–239. https://doi.org/10.1016/j.envsci.2007.07.004

Voß, J. P., & Simons, A. (2014). Instrument constituencies and the supply side of policy innovation: The social life of emissions trading. *Environmental Politics, 23*(5), 735–754.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

Epistemic Infrastructures: SDGs and the Making of Global Public Policy

Through an extensive discussion of the intertwinement of the multiple SDG policy arenas, actors and measurements, this book attempted to unpack the constitutive elements of the epistemic infrastructures of global public policy. The starting point of our analysis was the observation that, since knowledge and governance are closely interlinked (Jasanoff, 2004), an exploration of the epistemic base of any monitoring programme is central to an understanding of its structure and its governing effects over time. Conversely, an exploration of significant shifts in the production of global public policy requires a closer focus on the knowledge structures underpinning it. However, as the chapters of this book have shown, this is not merely a focus on 'evidence-making', which has been at the centre of the literature on public policy or the sociology of quantification thus far. On the contrary, this book has shown how, in the context of the SDGs, the production of global public policy must be understood as a complex interplay of material, techno-political and organisational structures within which statistical and governing knowledge is produced. In other words, policy is not simply informed or influenced by the numbers produced to guide it; instead, it is the spaces of the production of measurement themselves that have become arenas for the formation of global public policy. Therefore, the concept of 'epistemic infrastructure' offers a useful lens for understanding how these spaces come to be and how they become influential.

The following sections of this concluding chapter will theorise the first and second 'orders' of analysis of epistemic infrastructures: they will focus on the materialities of measurement, as well as the actors, processes and interlinkages

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*,
Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6_8

involved in the construction of the SDGs' epistemic infrastructure. Of course, there is no neat separation of the two 'orders' in reality: rather, we use it here as a heuristic, since it is on the entanglements of the material, social, political and organisational components of the infrastructure that the Sustainable Development Goals (SDGs) are built. Finally, the last section will discuss the formation of global public policy as a paradigmatic governing shift: it will show how and why global public policy is radically changing, as tensions and complementarities between technocratic and democratic imperatives are brought to the fore and become sites of struggle and formation of new ways of doing governing.

1 First-Order Level: The Material Underpinnings of the Infrastructure

As Chaps. 2, 3 and 4 of this book showed, the epistemic infrastructure of the SDGs is grounded in particular types of material building blocks: these are the data and the techniques of its collection, indicators and their categorisation into different tiers, harmonised data flows, narratives and stories, data visualisations, minutes of meetings and all other relevant inscriptions. Within the epistemic infrastructure, these different elements interact within one system—as it is a mix of technical, social and organisational materialities continuously interplaying (Bowker et al., 2009). Crucially, the infrastructure as a system does not emerge at once but rather it is a prolonged process of uneven development—and consequently, some material elements of the infrastructure will become obsolete or outdated before the infrastructure emerges as a whole (Star & Ruhleder, 1996).

These elements are the foundation of the infrastructure: they act as material representations of the phenomena at the centre of the SDGs—including poverty, education, health, tourism, migration and the whole range of policy arenas that 'sustainable development' brings together. The process of measurement occurs through inscription, which according to Latour are 'all the types of transformations through which an entity becomes materialised into a sign, an archive, a document, a piece of paper' (Latour, 1999, p. 306). Numbers, indicators and data within the epistemic infrastructure of the SDGs are such inscriptions: consequently, they are both material and semiotic—they are the physical manifestations of the areas of interest of the SDGs, but at the same time, they aim to reflect and also actively construct the meaning of these phenomena themselves. As such, as we have shown, they are not strictly representational—they are not merely constructions of the real world, but rather they are entities in their own right (Power, 2015).

The processes of constructing and validating indicators for the SDGs—alongside their classification into different 'tiers', depending on their technical qualities, the datasets that support them, but also political negotiations and compromises—are a case in point here. The SDG indicators are material and semiotic constructions because they become the space where the legitimacy and political acceptance of the exercise is constituted: as we carefully detailed, the Tier system works to classify indicators, based on whether there is adequate data and an acceptable methodology for measuring them. "Knowledge Production for the SDGs: Developing the Global Indicators" chapter discussed the ways that although negotiations for 'moving indicators up a tier' is apparently the aim, in fact, it is Tier II and III indicators—as the key 'placeholder numbers'—that enable and direct further action, even if still imprecise or incomplete. Consequently, Tier II and III indicators, as material-semiotic entities are not just failed numbers. Rather, they have significant generative power in terms of building the infrastructure: they are the subject of meetings (e.g., the High-Level Political Forums or Expert Group meetings), the focus of countless documents and analysis, and the reason for the development of networks, endless negotiations and consensus-building activities.

On the other hand, the book discussed the harmonisation of numbers as central to the process of *epistemic infrastructuring*. Although inherently requiring a specific social milieu, the process of commensuration is crucially a material practice that requires the numerical work of harmonisation in order to establish numbers that can govern (Espeland & Stevens, 1998; Bowker & Star, 2000; Timmermans & Epstein, 2010; Rottenburg et al., 2015). As "Harmonising Global Public Policy: Producing Global Standards, Local Data and Statistical Capacity Development" chapter showed, in the context of the SDGs, the material process of harmonising numbers is key to creating a unified field for global public policy. Along with the Inter-Agency and Expert Group for the Sustainable Development Goal Indicators' (IAEG-SDGs') work of harmonising methodologies, UN agencies have been given the responsibility for harmonising nationally produced data for the purposes of 'international comparability', to make available the means by which they 'produce and validate modelled estimates', and to coordinate with other International Organisations in order to verify such internationally comparable and sometimes imputed data (UNSD, 2017, p. 3). In order to compare social, economic, political or environmental conditions in geographically distinct locations, statisticians have to do the nitty-gritty, material work of harmonising data that may have been created with different methodologies, including different sampling techniques or different surveys. The book took a deep dive into explaining the process of producing metrics for the SDGs, in order to show the materiality and complexity of working with such diverse datasets to create knowledge for policy. Instead of glossing over these

processes as mere 'quantification', we showed how, at least in the field of the SDGs, disagreement, 'bad' numbers and discontent are not hidden and avoided but they are used as key sites for creating scientific and even democratic legitimacy.

Unlike physical infrastructures needing stability, the material underpinnings of the epistemic infrastructure are powerful because they are mobile (Latour, 1986). A central ingredient contributing to the fluidity and malleability of the epistemic infrastructure is narrative-making. Here, the book traced the entanglements of numerical, discursive and visual narratives in the production of stories that can persuade, move and bring actors together in constructing future utopias of a sustainable world with no hunger, no inequality and education for all. Through the analysis of narrative-making in the context of the SDGs, we showed how the SDG agenda is not exclusively a performance monitoring project, limited to requiring participating countries to regularly provide progress data towards the set goals. Rather, more ambitiously, the SDGs are presented in these documents not only as a 'governing' programme but rather a governing programme where values take the front stage. They aim to be seen as necessary, ethical, participatory and local; in fact, documents produced around the SDG often give little emphasis on the numbers themselves. In this sense, our findings align with previous research that argues that the SDGs are actively constructed as 'transformative' (Fischer & Fukuda-Parr, 2019, p. 376).

Discursively, the production of statistical data is often absent in the documents we analysed. Yet, this is precisely what the materiality of texts and narratives are about: the construction of an epistemic infrastructure that involves the creation of a knowledge system about ways of converting ideas about social life into numbers (Merry, 2016). The new orthodoxy of numbers does not need discussion or description; it has become a routinised 'way of doing things' and a way of enveloping the monitoring programme with meaning. Thus, the material work of narratives is essential to the political work of staging of all those necessary, yet often informal, frontstage and backstage rituals: the rhetoric, symbols and images required for transforming a deeply pragmatic and technocratic endeavour into a compelling 'story' for the audiences within and beyond its immediate cast and confines.

An important characteristic of the SDG 'story', however, is that it does not have a specific ending; instead, as is more common in modern storytelling, the 'story' offers several different endings that one can choose from. Indeed, the role of the material underpinnings of the SDGs (including the elaborate system of indicators, targets, report cards, custodian agencies, etc.) is not to strive for completeness but rather to keep the infrastructure open, incomplete and in constant movement. This point was made poignantly by Lampland (2010), who argued that the process of rationalisation of a different socio-political domain does not necessarily mean that

each stage is becoming more rational. Rather, there is a key role here to be played by the conditionality and transitionality of numbers. Incompleteness and an openness regarding the destination are motivating factors for keeping the measurement 'story' going.

Finally, the materiality of the infrastructure is invisible—it works when it becomes a taken-for-granted part of the background. As such, the infrastructure is visible when it breaks down—which is evident in cases where the contested SDG indicators and missing data are clearly visible in debates. As shown throughout the book, it was the indicators under debate and evaluation by custodian agencies and the IAEG-SDGs that produced extensive material in the form of open consultations, side meetings at the UN Statistical Commission or PowerPoint presentations. On the other hand, consensual measures are not visible—there is less material on them in these forms, and they simply become part of the reporting background. An important aspect of this (in)visibility of material components of the epistemic infrastructure is their missing parts—what is excluded from the infrastructure and why. This is what Bowker et al. (2009) referred to as the centrality of 'articulation work': the focus on what is missing and what is present in the infrastructure is not only a matter of (in)visibility but also a matter of strategic choice—or ignorance (McGoey, 2012). This is primarily the work of actors and networks, to whom we will turn next.

2 Second-Order Level: Interlinkages

At the second-order level of the epistemic infrastructure, 'rituals of verification' (Power, 1999) are conducted to link together the materialities of quantification—the indicators, the reports, the custodian agencies, the PowerPoint presentations and so on—into a web of relations, processes and practices. In Chaps. 5–7 of the book, we have shown how the concept of democracy has become central to these rituals of verification, how this network formation happens and how new forms of expert brokerage have become an engine for this formation. Presenting the constitution of the epistemic infrastructure at the second-order level in this way has allowed us to show how quantification in the SDG era has changed the way that different agencies in the global governance space—including United Nations agencies, member states, philanthropic organisations and civil society groups—engage with each other and produce a unified global public policy.

Drawing from actor-network theory (Callon, 1986) and the theory of epistemic communities (Haas, 1992), we argue that epistemic infrastructures are institutionalised through webs formed between human and non-human agents

and the knowledge practices that institutionalise such webs. Dynamic intra- and inter-organisational relationships (Fox, 2000) are central to the SDGs' global policy agenda. International organisations, member states and civil society groups and their quantification practices constitute an interdependent network for monitoring and producing globally agreed-upon goals, as we carefully outline in "SDGs and the Rise of an Epistemic Infrastructure: Actors' Networks, Partnerships and Conflicts in the Education SDG" chapter. As has been visible throughout the book, the epistemic communities producing governing knowledge for the SDGs have become quite expansive, while previous definitions of communities in the global space have focused on the role that elite communities with specialist knowledge have on influencing national policy (Haas, 1992). As can particularly be seen in "SDGs and the Politics of Reconciling the Dual Logic of Democracy and Technocracy" chapter, a wide range of stakeholders must be engaged in open consultations for refining and classifying indicators for inclusion in the monitoring framework—thereby constituting spaces of quantification both as privileged governing zones and founding the legitimacy of the SDGs as a global agenda on the ethos of democracy. These new epistemic communities have been structured explicitly to include previously excluded voices from member states from the Global South, civil society representatives and philanthropic and bilateral donor organisations. In the global governance space, knowledge does not travel from elite technical communities to policy spheres; instead, the production of knowledge by and through the SDG epistemic infrastructure is an iterative practice and always a negotiation between the technical and the political. Through the SDGs' particular form of network governance (Provan & Kenis, 2008), epistemic infrastructuring dedifferentiates these organisations and commensurates diverging social, political, economic and environmental phenomena through practices of data harmonisation. This is a continuous and often informal process, one of constant making and remaking of partnerships, connections and evidentiary claims.

Indeed, the notion of interdependency and collaboration amongst IOs was the starting point of the METRO investigation: in an increasingly interdependent, globalised world, IOs' work until few years ago remained surprisingly autonomous. The SDGs transformed this previous insularity of major IOs; the SDGs' governing architecture—with the sharing of responsibility between different IOs as the custodian agencies of the SDG indicators—led to a significant reconfiguration of how IOs produce metrics and brought them much closer together. It is at this level that we found that numbers—with their ability to simplify, stabilise and travel—reconfigure relationships, dependencies and structures of organisations and fields in fresh and politically salient ways. For International Organisations in particular, this has led to complex interdependencies, as they increasingly mobilise their resources

through their interaction with other IOs with comparable knowledge-producing abilities and interests: an IO's success may be seen as its power and influence over a larger regime of organisations that work towards specific policy directions, rather than through their insularity and autonomy (Raustiala & Victor, 2004).

Together with the requirement that IOs join forces in order to respond to global challenges came another demand: that of opening up the space of measurement as one that can and does facilitate democratic deliberation and engagement, especially for those previously excluded. These voices are representatives of countries, the civil society and especially actors from the Global South. As we explained in "SDGs and the Politics of Reconciling the Dual Logic of Democracy and Technocracy" chapter, this development was not a 'natural' and gradual development that followed on from previous work in the MDGs: rather, reconciliating measurement with inclusion of diverse and previously excluded voices was the flagship agenda of the SDGs, explicitly proclaimed and almost advertised as the key transformative change brought by the SDGs, as against the MDGs' ethos and governance. Marrying technical and political accountability is, however, not a straightforward task: the book discussed the tensions arising from the co-existence of the two, separately demarcated, types of accountability logic. As we have seen, striving for both political and scientific legitimacy has stirred uneasy discussions amongst the main players in the field. We found that actors were apprehensive about whether they can maintain the legitimacy of their expertise, when faced with the demand for technocracy to be politically accountable, too. On the other hand, those actors who strive for more inclusive and diverse decision-making do not want to be seen as less legitimate, or as compromising the validity and robustness of the data produced.

Finally, we also saw the emergence of a different type and style of expertise, one that does not only require a statistical nous, but is also dependent on achieving political mediation and consensus: "Expert Brokers: SDGs and the Emergence of New Forms of Expertise" chapter discussed in detail the rise of many International Organisations' experts as the key brokers ironing out discrepancies and disagreements both in scientific and in political terms. This kind of *brokering expertise*, walking on the tight rope of technical and political accountability, has been the fuel powering the epistemic infrastructure, constantly creating new spaces of expansion and growth (Bandola-Gill, 2022).

However, what is it about the SDGs that unravel the accountability conundrum and require such intense brokering activity for governing of the transnational? For a long time, the legitimacy of the production of quantitative knowledge related solely to its robustness and trustworthiness. Despite the emergence of more politically driven criticisms focusing on the (min)uses and effects of such monitoring

exercises, the debate on comparisons around country performance measurement, or the effects of these exercises on countries of the global South, tended to develop in a bifurcated way—with political and technical discussions neatly separated one from another, taking place in different spaces and relying on different parlances. Through proclaiming the SDGs as both a monitoring and democratic agenda, actors in this space have managed on the one hand to deal with the problem of power imbalances by creating a bottom-up narrative of development, whilst at the same time have diffused the problem by embracing it. The informality, multiplicity and dynamism of the epistemic and policy networks that bring the SDGs into being—as discussed in "SDGs and the Rise of an Epistemic Infrastructure: Actors' Networks, Partnerships and Conflicts in the Education SDG" chapter—are built precisely on the ability to simultaneously use different discourses and accountabilities, depending on the context and participant actors. Therefore, as we have shown in this book, the space of the meeting becomes the key venue where the complex technical statistical work of validating data for indicators takes place, but also the space where these data are presented to national representatives for their approval. Thus, the SDGs are not just a performance monitoring tool, like others: their transformative power lies in their success to be prescriptive, yet also appear as transparent, pluralistic, open and 'developmental'. In the end, it will matter little whether the goals are met or not—in having set up such a functioning, yet complex and fluid epistemic infrastructure, the SDGs are already successful in creating a unified global public policy and the means to build upon it.

To conclude, the epistemic infrastructure of the SDGs is built on a diverse set of actors and practices that 'harmonise' in a dual sense of the term—'infrastructuring' participatory governance *and* commensurating global public policy through the harmonisation of data production and indicator monitoring. It is certainly an open question whether this infrastructuring of participatory governance actually disrupts the power asymmetries that have long structured the relationships between UN agencies and countries in the Global North on the one hand, and countries in the Global South on the other. However, we argue that the epistemic infrastructure of the SDGs explicitly creates interdependencies between all these actors in the act of producing a common global public policy because of and despite these power differentials, making quantification the common denominator of global public policy in the process. The MDGs included eight goals that singled out development problems for member states in the Global South as conceived by organisations from the Global North, and they were focused much more on basic needs rather than promoting a platform for thriving. The SDGs produced a much more comprehensive global public policy space, and one which interpellated all countries as developing, thus producing a global governing paradigm. This is the third-order level of the epistemic infrastructure, discussed further in the next section.

3 Governing Numbers: Global Public Policy as a Paradigm Shift

The last half-century has seen immense social and political changes, such as the rise of globalisation, radical technological advancements and the shift from government to governance. Part and parcel of these changes has been the emergence of quantification as a way of knowing and governing highly dispersed, diverse and complex social realities. Indeed, quantification soon came to represent more than merely the measurement of people and practices. Rather, it emerged as the new way to govern: decision-making was not to be based on individual judgement, path dependency and a kind of 'connoisseurship' any longer. Instead, policymakers were encouraged to access new, evidence-based and international knowledge—or 'best practice', as it was often called. Different disciplines have richly analysed this radical change in the relationship of knowledge and policy and theorised on both its emancipatory but also distorting—and at times even destructive—effects. Quantitative knowledge—in all its manifestations, from data, to indicators and benchmarks, rankings and algorithms—came to be a key instrument in steering policy directions at all levels of government and often people's own personal lives; it became all-encompassing and, for some, a force of social transformation in itself.

Thus, what is it about the SDGs as a global monitoring programme that is any different from quantification as we already know it? A quick, superficial analysis of the SDGs would see them as following the line of work which began with the MDGs; indeed, many still see the SDGs as simply the renewed commitment of nations to development, using goal-setting as the key instrument to nudge countries into increased attention to areas such as education and poverty (e.g., Muchhala & Sengupta, 2014). Contrary to such accounts, we approached the SDGs as transformational: their potential for change has been outlined both in the official documents and the literature on sustainable development. Fukuda-Parr and McNeill (2019), for example, have persuasively written about the inclusion of Global South in the monitoring decision-making and process, as a key new feature in global governance. We build on this literature and move beyond it, to suggest that the SDGs are transforming the production of global public policy: through the construction of the epistemic infrastructure that gives them both substance and meaning, as well as a future direction, they are shaping a new governing paradigm *for* and *of* global public policy.

However, how has this come about? As the different chapters in this book have discussed, the SDGs have brought the two constituting elements of global

governance much closer together and in programmatic fashion: that is, on the one hand, the epistemological underpinnings that set the agenda of global public policy, and on the other, the governing architecture to produce it. By using the term 'epistemic' infrastructure (rather than measurement or knowledge infrastructure), our research highlighted the contribution of quantification to the rise of a broader epistemological agenda—one where quantification emerges as a cognitive scaffolding for thinking and understanding global problems by creating the conditions under which they become *knowable*. The arguments developed in this book showed how the notion of epistemic infrastructure has a double function: first, through bringing together discursive and numerical materialities, actors and networks, it captures larger epistemological questions; that is, the basic frames within which we understand the world and its challenges. Second, the epistemic infrastructure is instrumental in combining these new epistemologies with the structures and institutions that support, sustain and ultimately transform the measurement agenda into a decision-making space. Governing here is not just merely informed by numbers, or pursued by numbers. Instead, we see the SDGs as establishing a new governing paradigm, in that measurement (numbers) and governing are co-constructed: the measurement space is advocated and utilised as key space for achieving political consensus and shaping global public policy directions. New ideas emerge and bring new directions for envisioning and doing governing: interpretative flexibility, openness, (re)politicisation, reflexivity and democratisation are key discourses and proclaimed aims for the new governing paradigm that the SDGs represent. This is because the SDGs, analysed as an epistemic infrastructure, do not recast quantification merely as a tool in the arsenal of policy instrumentation and change. Instead, quantification is institutionalised as the very core of the governance of sustainable development.

Further, quantification in the context of the SDGs does not only represent the power of numbers to offer objective evaluations of the current state of the world. Although numbers have always been seen as powerful in their ability to offer neutral, a-political, fast and stable knowledge, giving legitimacy and authority to their producers and users alike, quantification within the SDGs represents a more fundamental shift than simply continuing to capitalise on the raw power of numbers to persuade. Instead, we see the proponents of quantification as gathering steam and building on the incremental gains achieved at the second half of the twentieth and early twenty-first century, in order to consolidate the ideas, norms, values and cultures of the epistemic turn in global governance. Their efforts have resulted in a deep shift that now renders quantification of the sole and most powerful infrastructural complex that has acquired the qualities and affordances of a macro-social policy paradigm (Hall, 1993). Stated otherwise, quantification in global governance

has consolidated previously disparate and even contradictory ideas into the making of a new *leitmotiv*, that can simultaneously be as complex as the vast datasets that feed it, or as simple as a new motto: 'leave no one behind'. Such metaphors and all their attendant elaborations structure the future: in the context of global public policy and the SDGs, quantification has become embedded in processes, decision-making, monitoring and accountability mechanisms that cannot be 'undone'. Decision-making in virtually all fields of global public policy takes place within the context of a particular set of ideas that recognise quantification as the governing frame that is more legitimate than others; as a result, quantification privileges, fosters and materialises some lines of policy direction over others.

Hence, we see the epistemic infrastructure of the SDGs as portraying a paradigmatic shift in the governance of global public policy. First, as this book has shown, the interdependencies of expert International Organisations, with country representatives, civil society, philanthropists and professional entities, have led to *re-politicisation of quantification*: the discourses of democratisation, participation, bottom-up country buy-in were a key fundamental change from the previous MDGs' era. The SDGs were from the start premised on a new, horizontal structure, where countries were considered centre stage; that is, countries are not merely participants, but are—theoretically, at least—in charge of the process. Crucially, this shift was the outcome of struggle and of the determination of countries of the Global South to change the narrative.[1]

Second, the SDGs acquired *global scope and reach*, as for the first time they framed development as not a requirement for the South only; instead, all countries are seen as continuously developing. Additionally, the scope of the agenda is much larger: goals have been set for a much larger spectrum of public policy arenas, even when there was no data to back up the vast majority of the goals put forward. More importantly, these policy arenas and the goals associated with them were seen as interlinked and as interacting with one another, giving rise to the emergence of a global public policy field centred around the notion of sustainability. Nonetheless, their translation into quantified targets soon gave rise to criticisms that some goals were contradictory and achieving one would be counterproductive for achieving another. Examples of such critique have primarily been expressed in relation to the economic growth goals contradicting the climate change or the global health ones (e.g., Hangoma & Surgey, 2019).

Third, as chapters in this book have shown, there is *interdependency and fragmentation* of the knowledge producers at the global stage. The range of producers

[1] Colombia's role in pushing for such a bottom-up governing structure has been notable and set the tone for a radically different path.

of statistical knowledge for the SDGs has grown substantially: IOs, civil society, NGOs, national statistical offices, donors and others. These actors are required to work together, whilst simultaneously maintaining their unique contribution and presence in the field. Such polar demands have generated a complex knowledge production arena, that simultaneously works together to standardise and integrate, whilst continuously growing the need for expanding the statistical capacity of nations. Despite the construction of global measures, these numerous 'statistical intermediaries' (Tichenor, 2022) create a sense of fragmentation and disjointedness, often necessary for the governing of such complex and fluid procedures.

Fourth, the SDGs are not an add-on to national policies; instead, they enter national agendas by being re-contextualised within national priorities and plans. Each country is expected to *nationalise* the SDGs—to produce a measurement infrastructure that matches as closely as possible to the global SDG framework to shape and be shaped by national priorities. One of the key modes of monitoring and enforcing this nationalisation is through Voluntary National Reviews, which are annually presented to the UN High-Level Political Forum.

Finally, and perhaps most importantly, the SDGs are not merely a performance monitoring agenda, sitting alongside the plentiful other global indexes and measurement tools. The goals have become pervasive in the everyday discourse of public and private sector marketing and corporate branding, in civil society and generally in a multitude of social institutions which want to be seen as adhering to the principles of sustainability and equality. It is precisely this influential positioning that sets it apart from previous endeavours of this kind.

To conclude, as the nature and breadth of transnational links and networks have expanded and the global diffusion of ideas, standards and policy practice has intensified (Stone, 2008), we observe new and changing geographies of policy (Peck & Theodore, 2015). Despite the complexity, fragmentation and instability of global public policy, sets of ideas can and do develop considerable coherence and persistence at the international level (Kennett, 2010). The analytical lens of an epistemic infrastructure allows us to not only explain these expanding connections but also propose an alternative theorisation of what global public policy is. As argued in this book, the theorisation of global public policy as an infrastructure goes beyond understanding policy in terms of its content. Global public policy is not just an assemblage of decisions affecting the global sphere; rather, it should be understood as a set of epistemic (infra)structures enabling forms of knowledge production, decision-making and interlinkages between actors. In that sense, global public policy is a process of *infrastructuring*—creating materialities, interlinkages between actors and common logics for action through which problems can be constructed and future policy directions can be drawn.

References

Bandola-Gill, J. (2022). Statistical entrepreneurs: the political work of infrastructuring the SDG indicators. Policy and Society. https://doi.org/10.1093/polsoc/puac013 (online first).

Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2009). Toward information infrastructure studies: Ways of knowing in a networked environment. In *International handbook of internet research* (pp. 97–117). Springer.

Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT Press.

Callon, M. (1986). The sociology of an actor-network: The case of the electric vehicle. In J. Law & A. Rip (Eds.), *Mapping the dynamics of science and technology*. Macmillan.

Espeland, W. N., & Stevens, M. L. (1998). Commensuration as a social process. *Annual Review of Sociology, 24*(1), 313–343.

Fox, S. (2000). Communities of practice, Foucault and Actor-network theory. *Journal of Management Studies, 37*(6), 853–868.

Fukuda-Parr, S. (2019). Keeping out extreme inequality from the SDG agenda: The politics of indicators. *Global Policy, 10*(Supplement 1), 61–69.

Fukuda-Parr, S., & McNeill, D. (2019). Knowledge and politics in setting and measuring the SDGs: Introduction to special issue. *Global Policy, 10*(Supplement 1), 5–15.

Grek, S. (2022). Infrastructuring numbers: SDG4 and the generative power of failing metrics. *Policy & Society* (online first).

Haas, P. (1992). Introduction: Epistemic communities and international policy coordination. *International Organization, 46*(1), 1–35.

Hall, P. (1993). Policy paradigms, social learning, and the state: The case of economic policymaking in Britain. *Comparative Politics, 25*(3), 275–296.

Hangoma, P., & Surgey, G. (2019). Contradictions within the SDGs: Are sin taxes for health improvement at odds with employment and economic growth in Zambia. *Globalization and Health, 15*, 82.

Jasanoff, S. (Ed.). (2004). *States of knowledge: The co-production of science and the social order*. Routledge.

Kennett, P. (2010). Global perspectives on governance. In S. P. Osborne (Ed.), *The new public governance? Emerging perspectives on the theory and practice of public governance* (pp. 19–35). Routledge.

Lampland, M. (2010). False numbers as formalizing practices. *Social Studies of Science, 40*(3), 377–404.

Latour, B. (1986). Visualization and cognition. *Knowledge and Society, 6*(6), 1–40.

Latour, B. (1999). Give me a laboratory and I will raise the world. In M. Biagioli (Ed.), *The science studies reader*. Routledge.

McGoey, L. (2012). Strategic unknowns: Towards a sociology of ignorance. *Economy and Society, 41*(1), 1–16.

Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. University of Chicago Press.

Muchhala, B., & Sengupta, M. (2014). A Déjà vu agenda or a development agenda? *Economic and Political Weekly, XLIX*(46), 28–30.

Peck, J., & Theodore, N. (2015). *Fast policy. Experimental statecraft at the threshold of neoliberalism*. University of Minnesota Press.

Power, M. (1999). *The audit society: Rituals of verification*. Oxford University Press.

Power, M. (2015). How accounting begins: Object formation and the accretion of infrastructure. *Accounting, Organizations and Society, 47*, 43–55.

Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory, 18*(2), 229–252.

Raustiala, K., & Victor, D. G. (2004). The regime complex for plant genetic resources. *International Organization, 58*(2), 277–309.

Rottenburg, R., Merry, S. E., Park, S.-J., & Mugler, J. (Eds.). (2015). *The world of indicators: The making of governmental knowledge through quantification*. Cambridge University Press.

Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research, 7*(1), 111–134.

Stone, D. (2008). Global public policy, transnational policy communities, and their networks. *Policy Studies Journal, 36*(1), 19–38.

Tichenor, M. (2022). Calculating capacities: Statistical capacity development and producing epistemic infrastructures. *Policy & Society* (online first).

Timmermans, S., & Epstein, S. (2010). A world of standards but not a standard world: Toward a sociology of standards and standardization. *Annual Review of Sociology, 36*(1), 69–89.

United Nations Statistical Division (UNSD). (2017). *Work of the statistical commission pertaining to the 2030 agenda for sustainable development A/RES/71/313*. United Nations General Assembly, 10 July. United Nations.

Index

J. Bandola-Gill et al., *Governing the Sustainable Development Goals*, Sustainable Development Goals Series,
https://doi.org/10.1007/978-3-031-03938-6