



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The New Production of Expert Knowledge

Education, Quantification and Utopia

Citation for published version:

Grek, S 2023, *The New Production of Expert Knowledge: Education, Quantification and Utopia*. Palgrave Studies in Science, Knowledge and Policy, 1 edn, Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-46606-9>

Digital Object Identifier (DOI):

[10.1007/978-3-031-46606-9](https://doi.org/10.1007/978-3-031-46606-9)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





PALGRAVE STUDIES IN
SCIENCE, KNOWLEDGE AND POLICY

The New Production of Expert Knowledge

Education, Quantification and Utopia

Sotiria Grek

OPEN ACCESS

palgrave
macmillan

Palgrave Studies in Science, Knowledge and Policy

Series Editors

Katherine Smith, University of Strathclyde, Glasgow, UK

Sotiria Grek, Centre for Science, Knowledge and Policy,
University of Edinburgh, Edinburgh, UK

The increasing production of knowledge for governance has resulted in major shifts in governance practices and public policy, especially for the work of knowledge producers and those identified as relevant ‘experts’. Given the centrality of knowledge to the governing of contemporary societies, how can we theorise the politics of knowledge for policy and governance and the strategies of policy influence?

Palgrave Studies in Knowledge, Policy and Governance invites social scientists to explore the construction of knowledge and expertise in governance processes across policy fields (e.g. crime, education, genomics, health, migration, sustainability, etc). We support interdisciplinary scholarship that deals with theorising real world issues, as well as practice-based perspectives. Contributors commonly work within Politics, Social Policy and Sociology but we welcome proposals from a wide range of disciplinary settings as well as those working in the policy world. The series encompasses diverse topics, methods and disciplines and we are keen to receive proposals for solo-authored, co-authored and edited books in either standard Palgrave or (shorter) Palgrave Pivot formats.

Please contact the series editors, Sotiria Grek (Sotiria.Grek@ed.ac.uk) and Kat Smith (Katherine.smith.100@strath.ac.uk) to discuss your initial ideas and outline proposals.

Kat Smith is Professor of Public Health Policy at the University of Strathclyde. Dr Sotiria Grek is Senior Lecturer at the University of Edinburgh. They are both members of SKAPE (the Centre for Science, Knowledge and Policy at Edinburgh).

Sotiria Grek

The New Production of Expert Knowledge

Education, Quantification and Utopia

palgrave
macmillan

Sotiria Grek
University of Edinburgh
Edinburgh, UK



ISSN 2947-5805 ISSN 2947-5813 (electronic)
Palgrave Studies in Science, Knowledge and Policy
ISBN 978-3-031-46605-2 ISBN 978-3-031-46606-9 (eBook)
<https://doi.org/10.1007/978-3-031-46606-9>

© The Editor(s) (if applicable) and The Author(s) 2024. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover credit: Filo

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

Paper in this product is recyclable.

For Asterios, Evie, Electra and Torcail
and
in memory of my father,
Yannis Grek,
who always dreamt of being a teacher and a poet.

WORDS OF THANKS

Amidst the uncertainties, puzzlements and all the ups and downs of being an academic researcher, there is one lesson that I have learnt well: research is a collective, rather than individual, endeavour. Although there are plenty of moments of personal questioning and introspection, good academic research is always the outcome of the coming together of many people, ideas, places and moments in time.

This book is no different. Many people have contributed directly or indirectly to its writing. First, I would like to thank the European Research Council for having generously funded the project ‘International Organisations and the Rise of a Global Metrological Field’ (METRO) (2017–2022, no. 715125), which was the main source of findings and analysis for this book. I would also like to thank the project postdoctoral research fellows for having worked so committedly, creatively and hard throughout the project, and particularly during the global pandemic of COVID-19 and beyond it: Drs. Justyna Bandola-Gill and Marlee Tichenor, thank you from the bottom of my heart. I have learned so much from you; the project would never have achieved all it did if it weren’t for your hard work, camaraderie and endless creativity. I offer my gratitude to Dr. Jennifer Ferguson who did METRO’s social network analysis. Many thanks to Drs. Matteo Ronzani and Ian Russell, METRO postdoctoral fellows, who also contributed greatly to the project at different stages of its development. Many thanks to the METRO’s visiting fellows, Dr. Marianna Zieleńska and Dr. Te-Anne Robles, as well as the project’s

administrators, Eirini Souri, Jack Thorburn and Anne Mourad. Above all, my warmest thanks go to all the project's research participants who so generously gave us their time and knowledge; the project depended on their goodwill and I will always feel indebted for their frankness and reflective wisdom.

The Department of Social Policy at the School of Social and Political Science, University of Edinburgh, has been my intellectual home for the last 12 years. Heartfelt thanks to all my wonderful colleagues who have supported my work and generously offered constructive feedback. In particular, I would like to thank Lindsay Paterson, Jeff Collin, Richard Freeman, Markus Ketola, Helen Packwood, Kaveri Qureshi and Rob Ralston. Beyond Social Policy, the Centre for Science, Knowledge and Policy (SKAPE) has always offered an informal and productive space for testing the ideas that shaped this book: many thanks go to the wonderful current and former SKAPE colleagues, and especially Steve Sturdy, Eugenia Rodriguez, Irvine Lapsley, Katherine Smith, Philip Cook and Marc Geddes. My colleagues at the 'Policy, Publics and Action' research group have been generous with their feedback and support. I owe special thanks to all previous and current doctoral colleagues Nikos Kanellopoulos, Natalie Papanastasiou, Emily Cameron-Blake, Hilary MacDonogh, Veronika Maricic, Mark Wong and Iryna Kushnir. My thanks to the wonderfully talented Anna Pultar; she is one of a kind and I have been so lucky to have met and worked with her. The last few months of writing were made easier by Claire Houghton, Amrit Virk, Sudeepa Abeyasinghe and Carin Runciman during our regular Friday writing retreats. I would like to thank Fiona McKay, John Devaney, Laura Jeffery and Janet Carsten for their continued friendship and support. Last but in no way least, my indebtedness goes to Christina Boswell: her scholarship, loyalty and friendship have been invaluable for more than a decade now.

Beyond Edinburgh, I have had the luck to be surrounded and supported by fantastic colleagues and friends from all over Europe and beyond. Many thanks to Christian Lundahl (Örebro), Joakim Landahl (Stockholm) and Linda Rönnerberg (Umeå); to Eric Mangez, Pieter Vanden Broeck (Louvain-la-Neuve) and Mathias Decuyper (KU Leuven); to Toni Verger, Clara Fontdevila and Alejandro Caravaca (Barcelona); and to Emiliano Grimaldi (Naples). I give thanks to Annabelle Littoz-Monnet (Geneva), Christian Ydesen (Aalborg), Romuald Normand (Strasbourg), Pieter Dahler-Larsen (Copenhagen),

Tobias Werron and Leopold Ringel (Bielefeld), Daniel Tröhler (Vienna), Nelli Piattoeva (Tampere), Gita Steiner-Khamsi (Columbia) and Florian Waldow (Berlin). I am indebted to Mary Morgan (LSE), Wendy Espeland (Northwestern) and Richard Rottenburg (Wits), all three of whom offered their time as keynote speakers at the METRO final conference and provided us with useful feedback as the work developed. A special mention to the late Sally Engle Merry (New York), who despite her ill health, offered her support and confidence in us. Special thanks to my good colleague and friend Paolo Landri (Naples) with whom I have spent more than four years editing the European Education Research Journal and have had numerous fruitful discussions over our common research interests. A big thanks to all the students and teachers of the SUSEES European summer school; I have loved working with you and I am appreciative of all your insightful comments and questions. A word of thanks to Bob Lingard, who has been an unwavering source of support and inspiration from the very start of my career. Finally, Jenny Ozga and Martin Lawn have always been great colleagues and close friends; from 2006 when I first met them to this day, they have always generously given me their guidance and advice. Every single person in this long list has contributed to the thinking and writing of this book in different ways; many thanks to each and every one of you.

In the end, I come home. I inherited my sense of curiosity from my parents, Yannis and Roula Grek, who have always inspired me to learn, live and love unconditionally. My brother, Apostolos, has always been by my side, a shining example of honesty, intelligence, musical flair and kindness. *Μαμά, Μπαμπά, Αποστόλη*, this book would never have been written without your unending love and encouragement. I would also like to thank Yannis, Dimitris and Vicky for their beautiful songs and smiles; I will always be your Lila. Many thanks to my best friend Maria, to Dimitris, to my goddaughter Athanassia and to Sofia for the wonderful summers in Halkidiki, in Rhodes and in Tinos—without your friendship, I could never have coped with the long Edinburgh winters. Closer to my adopted home, I would like to extend my gratitude to my friends Kalia, Natasa, Kristina, Nikos, Elpida, Maria, Farah and Eugenia. Many thanks to Norah, who supported me from my very first steps in the new country; without her, an academic career would have been beyond reach.

This book could never have been written without the love and support of my husband, Torcail, who came to me like an angel from the sky. His kindness, calmness, humour and companionship have gifted me a

beautiful life journey together; I look forward to all the roads still to be travelled.

Finally, I am indebted to my three wonderful children, whom I adore: Asterios Malcolm, whose warmth, intelligence and empathy are beyond measure; Evie Christina, full of artistic brilliance and love, and with a unique form of wisdom that I will always look up to; and Electra Ioanna, a caring, loving girl with so much curiosity, affection and spark. You are the absolute best and nothing is worth more than you three do.

As it happens, these last few words are being written on the kitchen table, at 14 Anakreontos Street, Harilaou, in my home city of Thessaloniki—the home I grew up in.

Everything has changed and all remains the same, and I feel enormously lucky for that.

CONTENTS

1	The New Production of Expert Knowledge in Education: An Overview	1
1	<i>Introduction</i>	1
2	<i>Global Numbers and the Work of IOs</i>	4
3	<i>Governing Knowledge, Experts and Data in the Twenty-First Century</i>	6
4	<i>From Mode 1 Mode 2 Science</i>	11
5	<i>The Book's Structure</i>	15
	<i>References</i>	18
2	Universality and Interdependence in Transnational Education Governance	23
1	<i>Introduction: Interdependence in a Complex World</i>	23
2	<i>'Governing by Numbers' in Transnational Governance</i>	25
3	<i>Theoretical Frame and Key Intermediary Concepts</i>	27
4	<i>The Case of the OECD as a Europeanising Actor: The Rise of the Interdependency of the OECD with the European Commission</i>	30
4.1	<i>International Comparative Assessments: the OECD's International Adult Literacy Survey (IALS) and PISA</i>	31

5	<i>From Dissensus to Universality: The Case of Narrative-Making in the SDG4</i>	38
5.1	<i>The 2014 Muscat Agreement</i>	39
5.2	<i>The 2015 Incheon Declaration</i>	45
6	<i>International Organisations: Interplay and Interdependence in the Making of the Global Metrological Field</i>	49
	<i>References</i>	54
3	The Rise of Mono-disciplinarity: Learning, Economics and the Production of Non-knowledge	61
1	<i>Introduction</i>	61
2	<i>The Contours of SDG4: Theorising with Epistemic Infrastructures</i>	62
3	<i>The Rise of Infrastructures: Vogue, Vague or ‘Really Useful Knowledge’?</i>	67
4	<i>From Schooling to Learning: The Incremental Building of an Infrastructural Base, 2006–2013</i>	69
5	<i>From IOs’ Advisory to Cooperative Role: Brokerage and Collaboration</i>	73
6	<i>Constructing Non-knowledge: Mono-disciplinarity and the Silencing of Alternative Perspectives</i>	75
7	<i>Discussion: Mono-disciplinarity and Quantification</i>	78
	<i>References</i>	84
4	Constructing Consensus by Data	89
1	<i>Introduction</i>	89
2	<i>The OECD Country Reviews of Education: The Case of Sweden</i>	90
2.1	<i>Socialisation and Learning in Governing: The OECD Reviews of National Policies for Education</i>	94
2.2	<i>The OECD Country Review of Sweden (2015) and the Foundation of the Swedish School Commission</i>	95
3	<i>UN’s ‘Participatory Turn’: Quantifying While Democratising</i>	99
4	<i>Experts or Brokers? UNESCO Institute of Statistics as a Trusted Actor</i>	104
4.1	<i>The Role of Meetings</i>	109
5	<i>Discussion</i>	112
	<i>References</i>	118

5	Beyond Objectivity? Storytelling and Reflexivity as Expert Work	123
1	<i>Introduction: Stories and Reflexivities in Global Education Governance</i>	123
2	<i>Expert Reflexivity: STS and the Work of Qualification</i>	125
3	<i>Playing God: Education Data Visualisations and the Art of World-Making</i>	129
3.1	<i>'No-One Left Behind'—Data Storytelling as Reflexivity-Making?</i>	130
4	<i>Instrumental Reflexivity and Expert Work</i>	136
5	<i>Experts' Reflexivity in Global Education Governance: The Role of Visuals and Stories</i>	140
	<i>References</i>	147
6	Navigating the Market of Measurement: Data, Quality, and Competition	153
1	<i>Introduction</i>	153
2	<i>The Market of Learning Assessments</i>	154
3	<i>The Case of the Quality Assurance Market in HE in Europe</i>	159
3.1	<i>Europeanisation as a Concept and as a Research Conundrum</i>	161
3.2	<i>Quality Assurance in Higher Education in Europe</i>	163
3.3	<i>Quality, Data and Standards</i>	165
3.4	<i>The Market of Actors</i>	171
4	<i>Discussion</i>	177
	<i>References</i>	178
7	Quantification and Utopia: New Forms of Expert Knowledge Production in Global Education Governance	183
1	<i>Introduction</i>	183
2	<i>From Mode 2 to the Production of New Expert Knowledge in the Twenty-first Century</i>	185
2.1	<i>Global Data for Universal Values</i>	186
2.2	<i>Monodisciplinary Visions of a Complex Education World</i>	188
2.3	<i>Perfect Brokers of Imperfect Numbers</i>	191
2.4	<i>Reflexive Experts of Datafied Systems</i>	192

2.5	<i>The Market of Measures</i>	194
3	<i>Education, Quantification and Utopia</i>	196
	<i>References</i>	202
	Index	207

LIST OF FIGURES

Chapter 5

Fig. 1	Front webpage of Left Behind visualisation	132
Fig. 2	Snapshots of Left Behind visualisation (01)	134
Fig. 3	Snapshots of Left Behind visualisation (02)	135
Fig. 4	Left Behind last page	136

Chapter 6

Fig. 1	Image from the World Education Blog ‘The Learning Assessment Market: pointers for countries’ (Montoya & Crouch, 2019b)	159
--------	--	-----

LIST OF TABLES

Chapter 1

Table 1	Attributes of Mode 1 and Mode 2 science (based on Gibbons et al., 1994)	12
Table 2	From Modes 1 and 2 to the new expert knowledge production	15

Chapter 3

Table 1	The SDG4 indicators are as follows: 4.1.1 on reading and maths proficiency; 4.2.1/2 on early childhood; 4.3.1 on VET; 4.4.1 on ICT skills; 4.5.1 on gender equality; 4.6.1 on adult literacy and numeracy; 4.7.1 on global citizenship and sustainable development. Available at http://uis.unesco.org/sites/default/files/documents/11-global-indicators-sdg4-cheat-sheet-2018-en.pdf	76
---------	--	----

Chapter 6

Table 1	The market of learning assessments (Fontdevila, 2023)	155
---------	---	-----



CHAPTER 1

The New Production of Expert Knowledge in Education: An Overview

1 INTRODUCTION

Over more than half a century, the dominance of International Organisations (IOs) in the production of global metrics has transformed global governance. However, amidst the avid critics and unapologetic fans of ‘governing by numbers’, it is still surprising that we know so little about the ways in which global processes of quantification are reconfiguring IOs’ work in the fast-moving field of global challenges. Metrics have infiltrated not only organisational cultures and the environments these organisations inhabit; crucially, they are reshaping the ways IOs co-exist, compete and survive in an increasingly datafied, yet uncertain world. Recent decades have seen fervent activity to build working collaborations and broad alliances for producing expert knowledge on global challenges. Financial investment in IOs’ collaborations is increasing and so is hope, as more and more policy actors place emphasis on global synergies and partnerships. Given the moral dimension that these measurements of progress have taken, as well as the enormous human and environmental cost of

their failures, *The New Production of Expert Knowledge: Education, Quantification and Utopia* casts light on the role of International Organisations in producing expert knowledge for transnational education governance.¹

Building on the findings of the European Research Council METRO research project (2017–2022), the book offers an interdisciplinary analysis of the interrelationships of International Organisations (IOs) in constructing the global metrological field. Education is the focal case for this examination: despite the prevalent idea that education is a predominantly national matter, IOs have been central to processes of standardisation, de-contextualisation and the performance management of education through numbers. As a result, they have been instrumental in commensurating, and therefore transforming the field. In addition, education governance has been attracting larger policy significance, as it is increasingly considered central to both economic prosperity and social cohesion. Education is a key element in the newly emergent well-being and ‘better life’ strategies that have prevailed in the statistical governing project post the 2009 financial crisis (Stiglitz et al., 2009); more recently, after the global pandemic of 2020–2023, it is again education and its ‘losses’ that are counted and planned for. Last but not least, education is a key policy arena in monitoring projects, given first, its history in the establishment of testing and the associated calculative technologies that came with it, as well as its congruence with efforts to use ‘softer’ data sets for calculating the social. These are just some of the many reasons that large IOs like UNESCO, the OECD, the European Commission and the World Bank have invested large amounts of data and expertise in education from the mid-twentieth century on. The encoding of education data processes and organisational cultures that these commensurating practices require (in order for data to be shared and co-produced) represents a microcosm of the workings of quantification for transnational governance. In other words, the examination of the production of new governing knowledge in education by IOs is a unique opportunity to open, rather than stack yet another ‘black box’ in the field of global governance (Bhuta, 2012).

However, the book goes beyond an examination of the role of IOs in ‘governing by numbers’. As previously suggested, multiple bodies of

¹ Here we follow Djelic and Sahlin-Andersson’s preference of the term ‘transnational’ versus ‘global’ governance, since ‘the label “transnational” suggests entanglement and blurred boundaries to a degree that the term “global” could not’ (2006, p. 4—for a more developed argument see also Hannerz, 1996).

knowledge are brought together in order to analyse the role metrics play in reshaping the relationships between the data collectors themselves. Thus, a core objective of the analysis is to understand the role of quantification in the interplay of IOs. Second, although there have been some in-depth studies of the impact of measurement on reforms in various policy fields, little attention has been paid at those early, yet crucial, venues, actors and activities that determine processes of problematisation (the construction of the ‘problem’) and institutionalisation (the moment the ‘problem’ enters institutional agendas). Third, and most important, this book’s starting point is that numbers and (international) organisations have come to be mutually constitutive. Numbers move: this seemingly simple, yet unique quality has created a fluidity between internal organisational arrangements and external environments, as well as among IOs themselves. Hence, going beyond classic organisational sociology’s distinction between internal structures and external contingencies and environments, this book discusses the ways that numbers—with their qualities to simplify, stabilise and travel—reconfigure relationships, dependencies and structures of organisations and fields in fresh and politically salient ways; in other words, they come to govern them.

As previously suggested, the book analyses in-depth empirical findings, as they emerged from the European Research Council funded project ‘International Organisations and the Rise of a Global Metrological Field’ (METRO, for short). The project run in the period between 2017 and 2022 and involved a comparative case study of different policy fields, examining the Sustainable Development Goals as a whole, but also focusing deeper on the cases of education (in particular, SDG4 and the emergence of the ‘European Education Area’), the global monitoring of poverty and the case of statistical capacity development (as it became increasingly key across all the SDGs). The study used mixed methods, combining textual analysis, social network analysis, as well as over 80 interviews with key experts in International Organisations, including the World Bank; the United Nations Educational, Scientific and Cultural Organization (UNESCO); the UN Children’s Fund (UNICEF); the UN Development Programme (UNDP); the World Health Organization (WHO); the UN Statistical Division (UNSD); and the Partnership in Statistics for Development in the 21st Century (PARIS21). In addition, the research team drew on the careful analysis of official documents produced by this epistemic community, 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. Empirically, the central analytical approach was inspired by grounded theory, entailing multiple rounds of coding (including descriptive, focused and theoretical coding) (Charmaz, 2006). Conceptually, the project built on and synthesised political sociology, Science and Technology Studies (STS) and theoretical strands from the field of the social studies of metrics.

2 GLOBAL NUMBERS AND THE WORK OF IOS

Despite the renewed prominence that is given to the need for alliance-building by IOs, collaboration has always been central to their operation, since they have traditionally needed to work closely with governments, NGOs and the private sector. Yet, the complexity of ‘wicked’ problems, ‘donor duplication’ (Ringel-Bickelmeier & Ringel, 2010), resource-pooling and data overload have become some of the most common reasons that IOs are increasingly compelled to work together. Indeed, most major global strategies, such as the Millennium Development Goals (2000–2015), the post-2015 Sustainable Development Goals or major education testing regimes, such as the OECD Programme for International Student Assessment (PISA), are collaborative endeavours, dependent on pooling of resources and expertise. How do these IOs learn from one another? In the making of numbers, how do they negotiate knowledge controversies and share the expertise they generate? How do they actively produce collective sense-making (Weick, 1995) and issue-framing strategies (Baumgartner & Jones, 1993)? How much do we know about their expert networks? How do they manage to produce expertise together, while maintaining their unique branding and contribution in the field? Ultimately, if rating and ranking practices are a ‘zero-sum’ game for the assessed, how much do we know about the rules of the game for the assessors?

Although questions still abound, at least in the field of education, the production of data and numbers has—for some time now—become a key mechanism of both knowing and governing the field. Complex statistical systems—best exemplified by the Sustainable Development Goals (SDGs) introduced in 2015—as well as performance measurement instruments, such as the OECD’s Programme for International Student Assessment (PISA), have emerged as tools for both monitoring and steering

global education action. This process of hyper-quantification has had far-reaching consequences: as numbers evolve from merely instruments for governance to ‘civic epistemologies’ (Jasanoff, 2004); they reshape the broader context in which knowledge about problems is produced, and thus change political identities, relationships and institutions (Bandola-Gill et al., 2022; Miller, 2001). At the same time, quantification is as powerful as it can be paradoxical: measurement is not a neutral activity but located at the intersection of diverse (and often competing) epistemic and value orders.

More specifically, we know well by now that the power of quantification is firmly positioned in such epistemic virtues as objectivity and political neutrality (Porter, 1995). Historically, the power of numbers stemmed from their ability to represent—and construct—governing problems, underpinned by the technocratic legitimacy of the seemingly apolitical statistical method (Grek, 2010). Nevertheless, one of the core arguments of this book is that the power of numbers is equally grounded in their political value—and this value is increasingly foregrounded on the global arena. Thus, the former push for depoliticizing decision-making on the basis of evidence has been more recently counterbalanced by the *re-politicisation* of education metrics, particularly as a result of the ‘participatory turn’ in the global monitoring systems (Bandola-Gill et al., 2022): the ‘turn’ necessitated the wider participation of actors in number-making, including from countries from the Global South (Fukuda-Parr & McNeill, 2019), with the aspiration to create opportunities for more democratised statistical systems (Milan & Treré, 2019; Tichenor, 2022). Increasingly, the production of numbers is expected to go beyond ‘global’ numbers and instead to account for ‘local’ politics and needs—or at least to give them equal weight, in that no global numbers can be produced without the active co-option of local actors and their needs.

This tension between technical and political accountability, or in other words, between authoritative and democratising numbers, is at the crux of this book: although the empirical work behind the METRO project set off from an interest in exploring the interplay between IOs, what the field-work manifestly showed is that the notion of interplay goes well beyond the confines of understanding how IOs collaborate in the production of metrics. Instead, the research team was confronted by experts in the field who did not see them as authoritative actors, distant from their field of enquiry in producing objective knowledge about it; on the contrary, they saw themselves as caring figures, that accepted the political power of

numbers, and were largely dependent—restrained and enabled in equal measure—on the complexities of the participatory processes of producing global knowledge and governing in a post-truth world.

Before going into the detail of the main arguments that are presented in this book, I would first like to offer a brief overview of the development of expertise in transnational education governance from the late twentieth century today: what began as the first efforts to establish comparative data in education, swiftly turned into evidence-based policy and ‘what works’ in the 1990s, to develop further to the establishment of European and global soft governance tools, such as the Lisbon Agenda at the European level and the Millennium Development Goals globally. One of the core STS conceptualisations on the shift from Mode 1 to Mode 2 scientific knowledge production will be briefly discussed here, as a useful tool for explaining the many further developments and changes in the making of global education expertise. While I fully acknowledge the limitations of the Mode 1 to Mode 2 frame, I use it here as a heuristic device that allows to present both the data and the analysis more clearly. Finally, I will introduce the book’s chapters, with the aim to show the ways that post-2015 education expertise is characterised by new qualities and characteristics in the ways that it is produced, negotiated and communicated.

3 GOVERNING KNOWLEDGE, EXPERTS AND DATA IN THE TWENTY-FIRST CENTURY

The last thirty years have seen a major shift in the production of education research for policy. In this section, I examine the specific case history of the emergence of this new knowledge production regime in the field of education research, starting from Europe and the United States and spreading globally. During this period, changes in policymaking, which have been summarised as a shift from government to governance, and changes in knowledge production (including increasingly algorithmic knowledge and artificial intelligence) (Gibbons et al., 1994; Nowotny et al., 2001) have come together symbiotically: changing governance processes and norms create the conditions for new kinds of knowledge production, and such production of expert knowledge for policy becomes a key resource for governing the ‘perma-crises’ that contemporary societies find themselves in, experiencing the compounding challenges of climate change, global pandemics, inequalities and more.

The idea of a ‘governance turn’, as marking a significant shift in governing practices in Europe and beyond, continues to be of relevance in the analysis of education research (Beukel, 2001; Hooghe & Marks, 2001; Mayntz, 1994). In brief, governance describes a move from centralised and vertical hierarchical forms of regulation to decentralised, horizontal and networked forms: this is a phenomenon claimed by some to be global (Rosenau, 1999) though this is hotly disputed, both pre- and post the recent global pandemic. However, whatever the extent of variation, governance is described in ways that reflect broad patterns that themselves may be understood to discursively reflect dominant political forces. The increasing involvement of private actors in the production of knowledge in education has multiplied these effects, since the emergent ‘stakeholderisation’ of global governance has led to diminished rather than increased democratic decision-making.

As the chapters in this book will discuss, novel governing practices promote ways of controlling and shaping behaviour (Hood et al., 2001) that mix material and discursive strategies: the discursive mobilisation of new norms and values is combined with external regulatory mechanisms (for example, competitive indicators of performance or global monitoring regimes, such as the SDGs) which together seek to transform the conduct of organisations and individuals. As a result, transnational governance is produced through the construction of ever-evolving epistemic infrastructures where the technical and the political have become a single entangled mix: as will be discussed in the chapters of this book, in the name of the democratisation of knowledge production, the so-called pluralism of voices has paradoxically led to the further strengthening of monodisciplinary and datafied knowledge for governance (Grek, 2022; Tichenor et al., 2022).

In many European countries, we can trace a process of circulation of these discursive norms from the 1980s, and the simultaneous development of new regulatory forms: we observe deregulation accompanied by tighter specification (for example, in the field of education, the emergence of centrally prescribed curricula and testing regimes), the growth of technical accountability, and a dominance of new public management principles applied to the public sector. In education in particular, there has been a steady growth of governing through performance management around principles of decentralisation, devolution and deregulation as key principles of system restructuring (Whitty et al., 1998). Those key principles were not challenged—indeed, in many cases, they were

reinforced—by shifts in political parties in power in most European countries, and indeed in the governance of the European project itself, as new ‘imaginaries’ (Jessop, 2008) connected education closely with the rise of knowledge economies for improving growth and social cohesion (Mulderigg, 2008).

In the same timeframe, we can also chart the emergence of apparently new forms of knowledge that provide useful support for agendas that stress collaborative solutions and rapid adaptation, or that express ‘new institutional compatibilities’ (Nowotny et al., 2001), between knowledge production and use. In the era of neoliberalism, knowledge became internal, i.e. part of, rather than external to and distinct from the economic process. Economic growth was seen as dependent on maximising the outputs of knowledge workers and the productivity of knowledge resources. National education systems sought to ensure competitive advantage through the commercial exploitation and application of the knowledge produced by ‘research-intensive’ universities. Technologies enabled the instantaneous exchange of information. These exchanges transcend national boundaries, so the constraints of national economies give way to an interdependent global economy; the recent pandemic has accelerated the process of digitisation of education further. The funding, organisation and assessment of research quality are all affected by these developments. Kenway et al. (2004) illustrate the trend towards prioritising techno-scientific research and its modes of operation and organisation, so that research is increasingly concentrated in designated centres of excellence, organised in teams and characterised by differences in conditions of work and employment rights. Traditional intellectual autonomy is challenged by the need to meet industry needs and, as a consequence, science is becoming less a public good than more of a tradable commodity.

The centrality of research and knowledge production for growth helps to explain the enhanced research steering policy agendas across different national settings in Europe and beyond. Research steering processes emerge at the national level that promote particular methodologies and particular forms of measurement of research quality and recognition (for example various forms of metrics, benchmarking and citation indices). In addition, knowledge has further been commodified, through the emergence of a large data production and information industries, which is described by policymakers as promising greater transparency and hence quality for the public services, education included (Ball, 2007). These

trends reflect a perspective on education research that prioritises its ‘use-value’ and its problem-solving potential for policymakers, as key indicators of quality.

This increased significance of knowledge means that in the developed world, information and expertise have—for some time now—been more widely available and more widely distributed than ever before. At the same time, new governance forms promote the idea of transparency, public accountability, sustainability and the democratisation of knowledge as part of their strategic positioning; decoloniality has become the rallying cry of those who fight against the continued epistemic injustices of a knowledge system organised and decided upon by the Global North. Knowledge is drawn into supporting the legitimacy and authority of the social and political processes of new governance agendas. Discursively, knowledge and policy are produced as a form of cultural political economy (Jessop, 2008) which combines semiotic and material elements in changing the nature of knowledge production and its role in governing. Policymakers suggest that social cohesion and effective government now depend on integrating knowledge in decision-making processes. This positioning promotes an agenda for the future in which potentially disruptive energies (the rise of artificial intelligence, for example) are harnessed to promote a discourse of entrepreneurship and continuous scientific and technical advancements (Mulderigg, 2008, p. 167). As Bauman (1992) put it three decades ago, in a decentred, information-rich society, governance needs to use ‘science’ more actively to minimise risk, or—at the very least—to minimise anxiety about risk.

The production of expert knowledge in education is subjected to the same forms of regulation and risk management. As with other expert knowledge, it is applied, scientised knowledge, packaged in flows of data and tables. Knowledge production is equated with particular forms of data collection and comparison and its quality is judged in relation to its usefulness in assessing comparative performance. This transformation of the field of education is happening through the reshaping of the old institutions of schooling and post-compulsory education and their replacement with designs for (lifelong) learning, that require new, accessible and portable qualifications frameworks (Grek, 2008; Grek & Russell, 2023; Ozga et al., 2006) and through the development of new attitudes that instil responsibility and commitment to continuous self-improvement for schools and learners alike. The task of governing knowledge is to map and loosely link a complex space of flows of international and national actors

and data, with the aim of imposing its logic over scattered, segmented places or what Martin Lawn previously called ‘systemless’ systems (2013): in other words, the disarticulation of a public education system into political, spatial, contextual and increasingly commercial parts is loosely connected via data.

In most education systems around the world, systems of performance and quality management have learnt to provide ‘proof’ of the quality of their ‘outputs’. Middle class parents became experts in decoding and using this information, while policymakers are more dependent on and subject to the judgements of experts. Quality management regimes or the various systems of research quality assessment are input–output machines that contain team rules, rules of evaluation, cooperation and innovation. Experts propagate ‘efficiency myths’ that allow for the growth of quality management and professionals, including the teaching profession and academic researchers are reformed as active protagonists of quality systems. In this process, we see the ‘transformation’ rather than the transfer of knowledge, with the key element of scientific knowledge production—uncertainty—simultaneously (or, better, temporarily) removed and strengthened so as to necessitate the continued need for the further production of data (Stehr, 1994). The elimination of doubt and continuous affirmation of usefulness and ‘social impact’ are constructed discursively through the language of research assessment. Experts are ‘chosen’ for their capacity to provide what they often see or translate into technical advice and—increasingly—the presence of experts in such ‘user-driven’ activities counts as a quality indicator in itself (Lawn & Lingard, 2002).

To conclude, quantification and the production of expertise more broadly do not merely inform but have come to constitute a ‘state optic for governing’ (Scott, 1998). There are intimate and interwoven relationships between the development of state administrative structures, characterised by Latour (1987) as ‘centres of calculation’, and the development of standardisation, methodological approaches, technologies and related cognitive schemes of statistics and scientific thinking (Desrosières, 1998; Porter, 1995). This analysis is, of course rather at odds with the collaborative and socially embedded possibilities of co-production of knowledge as presented by Nowotny and her colleagues (2001, 2003), and runs against more recent developments around the decolonisation and democratisation of knowledge production more generally and expert knowledge for policy production in particular. The next section will briefly

discuss Mode 2 knowledge production, before moving to a description of the book's chapters and the ways the Mode 2 regime has further transformed as a result of the evolution of quantification.

4 FROM MODE 1 MODE 2 SCIENCE

The literature on new modes of knowledge production gained traction in the 1990s and 2000s as global challenges began emerging and presenting multiple elements of complexity and intertwinement (Crowley & Head, 2017). Therefore, it quickly became obvious (to some, at least) that new kinds of expert knowledge needed to be produced in order to deal with those global social transformations (Eyal, 2019). The central rationale underpinning this body of work in general (Etzkowitz & Leydesdorff, 2000; Funtowicz & Ravetz, 1993; Wesselink & Hoppe, 2011) and the concept of Mode 2 (Gibbons et al., 1994; Nowotny et al., 2001) in particular, was an observation that a change in societal values and practices inevitably results in an evolution of the epistemic structures of such society. Therefore, scholars working in the field of Science and Technology Studies understood transformations in governing—and particularly the shift from government to governance—as leading to changes to the systems of knowledge production (Miller & Rose, 2008).

As a result, *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies* published by Michael Gibbons and colleagues in 1994 argues that a new model of knowledge production has emerged in modern societies (a summary is presented in Table 1). According to Gibbons et al. (1994), the increasing complexity of societal issues had posed a challenge to academic knowledge production, leading to the replacement of traditional science (Mode 1) with new knowledge production (Mode 2).

First, in contrast to older paradigms and practices, knowledge production in the late twentieth century was seen as deeply embedded in society, rather than limited to universities: knowledge, it was argued, could not be solely produced in ivory towers any longer but had to take the 'context of application' into account. What STS scholars called 'contextualised science' assumed a deep embeddedness of science within society, whereby society in turn 'speaks back' to science (Gibbons et al., 1994, p. 50). Second, the new knowledge was seen as produced outside of the traditional disciplinary boundaries, as the new challenges require collaborations between experts with diverse disciplinary and institutional

Table 1 Attributes of Mode 1 and Mode 2 science (based on Gibbons et al., 1994)

<i>Mode 1</i>	<i>Mode 2</i>
Problems of knowledge are set and solved in a context governed by academic interests of a specific community	Knowledge is produced and carried out in a context of application
Based on the disciplines	Cross/Trans-disciplinary
Homogeneity	Heterogeneity
Hierarchical structure, and tends to preserve its form	Heterarchical and transient
Quality control through peer review	Socially accountable and reflexive

backgrounds (Gibbons et al., 1994). Consequently, both the empirical and the theoretical structure of knowledge produced within the Mode 2 differs from the traditional, siloed and disciplinary structures. Third, knowledge production went beyond the traditional academic structures and was made by a variety of actors, including government agencies, research centres, think tanks, international organisations and others. Fourth, the new knowledge production was seen as more responsive and reflexive to societal needs; the supremacy of social problems, versus the notion of scientific autonomy as the central cultural value of science, is a defining feature of science for society rather than the production of science for science's sake. Lastly, in the Mode 2 transformation, the quality of knowledge production is assessed by using the broader criteria of social, economic and cultural usability of the produced knowledge (Gibbons et al., 1994).

The concept of Mode 2 science gained considerable traction and moved beyond the production of scientific knowledge to also cover policy knowledge (Logar, 2011; Lövbrand, 2011; Raftery et al., 2016). Despite its intuitive appeal for explaining the changes in the production of knowledge for policy (Yearley, 2005), the work of Gibbons et al. (1994) sparked a broad debate. The Mode 2 concept was criticised for the lack of theoretical grounding (to which the authors responded in the second book—Nowotny et al., 2001), as well as an ahistorical view of the evolution of science. Thus, the critique of Mode 2 science highlighted the fact that the notion of science devoid of any practical consideration—the central quality of Mode 1—is very rare when seen from the historical perspective (for example Etzkowitz & Leydesdorff, 2000; Pestre, 2003).

However, the main criticism against Mode 2 revolved around what was perceived as an evolutionary perspective on knowledge-making, hence ignoring the political and social shaping of knowledge production in contemporary societies (Pestre, 2003). As Dominique Pestre points out, analysis of the discourse used to characterise Mode 2 knowledge (for example social relevance, responsibility, reflexivity, fluidity) largely eliminated alternative definitions of knowledge and knowledge production and created a strong normative pressure on researchers to enhance their responsiveness and usefulness. As he put it, the Mode 2 discourse conveys an ‘overly optimistic’ vision of the changes affecting science and society today. He goes on to comment that:

The authors may have underestimated the extent to which these transformations have been the result of political and social choices. This would mean recognising that the developments they describe are not cases of natural evolution, which have simply to be identified and acknowledged, but are, rather, articulated with alternative and conflicting social, economic and political projects. (Pestre, 2003, p. 246, emphasis in original)

As well as reminding us of the ways in which knowledge has always mattered to states and economic elites, Pestre underlines the importance of knowledge as a resource for changing social ideologies (ibid:250). The transformation of knowledge is linked to the transformation of capitalism in this analysis, showing how knowledge has both mirrored that shift and made it possible, thus creating new levels of interdependence, of the kind illustrated earlier in the discussion of performance data. This interdependence is also neatly captured in Nigel Thrift’s book *Knowing Capitalism* (Thrift, 2005), which illustrates how the cultural circuit of capitalism produces knowledge about itself and illuminates how capitalism has become knowledgeable and thus increasingly impinges on traditional academic preserves (Thrift, 2005, p. 21). Part of this process, Thrift argues, involves capital and traditional knowledge producers in the academy coming to ‘think more alike about thinking’ (ibid., p. 21).

Indeed, as the chapters of this book will discuss, the METRO research has shown that the production of expert knowledge for policy, at least at the transnational realm, has increased the *universality* and often *uniformity* of choices and outcomes of education policy-making, often leading to a much closer alignment of education with the economy. As a result, one of the key functions of experts is the *brokerage* of knowledge, in order

to find ways to create consensus and—paradoxically—transform the technocratic spaces of number-making to the ones that will also address the democratic deficit that such processes have been blamed for in the past. Thus, rather than the Mode 2 proclamation of hybridity and diversification of knowledge production, what we observe is increased universality of the policy agenda, as well as the technicisation of political issues through the transformation of the spaces of measurement into spaces of negotiation and political consensus-making.

Further, as we will see, rather than choosing trans- or even interdisciplinarity, the *mono-disciplinarity* of the dominance of economics has prevailed in the field not only of education research but also knowledge production for policy more broadly. In addition, despite Nowotny's argument about Mode 2 knowledge moving away from hierarchy of knowledge-making as pushed on by the state or the markets, we observe that *new markets of measurement* and indicators have emerged: the use and predominance of certain measurements over others determine their popularity and lead policymakers to react to them differently and often frenetically (as the PISA experience widely showed). In addition, as we have seen, not only do 'governance' and 'mode 2 knowledge' share a repertoire of defining terms, but they have also worked discursively to create images of progress and *democratisation*, to support inclusion, and to push for the *co-option of knowledge in governance*, dissolving the boundaries between them. The so-called governance turn is often defined in terms that echo this supposed transformation of knowledge from elitist to more democratic: a shift from centralised and vertical hierarchical forms of regulation to decentralised, horizontal, networked forms. Yet, thirty years on, rather than representing the potential for democratisation of either knowledge or governance, these forms have come to closely resemble the networking practices and open communications systems of global capital. Indeed, 'edu-businesses' are now also a key knowledge actor and a significant player in education systems around the world, especially after the global pandemic. Education research for policy often reflects the processes and instruments of knowledgeable capitalism and its dominant 'economic imaginaries' have established new organisational and numerical forms that have 'a performative, constitutive force' (Jessop, 2008:18).

Taking all these serious and valid criticisms of the Mode 1 to Mode 2 schema of knowledge transformation, as well as more recent moves to discuss a Mode 3 paradigm (which further legitimise and reinforce the

Table 2 From Modes 1 and 2 to the new expert knowledge production

<i>Mode 1 knowledge</i>	<i>Mode 2 knowledge</i>	<i>Expert knowledge</i>
University context	Context of application	Global/universal level
Disciplinarity	Transdisciplinarity	Post-disciplinarity/ Mono-disciplinarity (Economisation)
Homogeneity	Heterogeneity and Organisational Diversity	Brokerage/ consensus/ mediation
Autonomy	Social accountability and user reflexivity	Datafied accountability and expert reflexivity
Peer-review quality control	Extended quality control	The market of measures

critique against the normative character of Mode 2—see, for example, Carayannis & Campbell, 2011), *The New Production of Expert Knowledge: Education, Quantification and Utopia* utilises the schematic representation of Mode 2 to discuss expert knowledge production in the first quarter of the twenty-first century (Table 2).

The next and last section of this introductory chapter will outline the chapters in this book and explain how they cast light on the new expert knowledge-making.

5 THE BOOK'S STRUCTURE

Chapter 2 will discuss the move away from the specificity of the context of application to the universality and interdependence of global education metrics: as I shall show, instead of the production of contextual knowledge, quantification in transnational governance has led to the production of expertise that is thoroughly standardised, de-contextualised, interdependent and even universal. The chapter focuses on two empirical examples of international organisations that saw their status as knowledge producers and expert brokers rise over the last 20 years: these are the OECD, and its collaboration with the European Commission, as well as the UNESCO Institute of Statistics with its coordination of the SDG4. Through an analytical account of these organisations' key measurement

exercises, the chapter charts two key developments towards the production of decontextualised governing knowledge: these are the rise of the interdependence of IOs in the production of expertise; and secondly, the production of universal narratives of education progress and unity.

Chapter 3 analyses the ways that the production of data for education over the last three decades, despite the complexity and interdependency of policy problems in education, has not been interdisciplinary, but the opposite: it has primarily been dependent on the discipline of economics and the ensuing economisation of education policy as the preferred mode of producing knowledge for governing. The chapter mobilises relevant literature and uses empirical examples in order to offer two propositions: first, that instead of disciplinarity, global education governance is primarily dependent on a monodisciplinary knowledge production orientation; second, that an investigation of metrological realism needs to focus on the social construction of non-knowledge as a vital component of studying the epistemic authority of transnational institutions.

Chapter 4 turns the lens to the processes that influence and steer the production of expert knowledge in the global governance of education over the last 50 years. The chapter adopts the position that its construction is not ‘organic’—the product of traditional knowledge-making as it became dominant from the Enlightenment onwards—but rather the outcome of complex undertakings that often imbricate a wide variety of actors—both national and international, including decision-makers—and different fields. The chapter builds on the shift from Mode 1 to Mode 2 knowledge production (Gibbons et al., 1994) in order to document further changes to how expert knowledge is produced today: it argues that, at least in the field of global education governance, we see concerted efforts to produce expert knowledge that focuses equally on technocratic and political accountability, and that sees brokerage and consensus-making as the ultimate goals in an increasingly polarised and uncertain post-pandemic world.

Chapter 5 focuses on an analysis of the role of storytelling and reflexivity in further strengthening and legitimising quantification in global education governance. It examines two specific empirical examples that show, first, how data visualisations in education and sustainable development are changing in order to accommodate the construction of a more democratic and inclusive governing space; second, how policy and expert actors themselves use reflexivity as a way not only to understand and think about their daily policy work, but also to create spaces of alignment and

consensus. In that way, both storytelling and reflexivity can be seen as working instrumentally, enhancing and further embedding the work of ‘governing by numbers’, rather than displacing them.

Chapter 6 discusses the rise of the competition over measurement which has been structuring the relationships between IOs. The production of data to support comparative assessment and evaluation is one of IOs’ key organisational remits, therefore, they have vested interests in promoting the implementation of their measures over those of others. Consequently, what we observe in the global governance of education is not merely ‘governing by numbers’, but rather a navigation of the market of measurement; this can often lead to conflicts and controversies over statistical data collection, as well as new partnerships and collaborations. Thus, it becomes obvious that it is not merely epistemic authority that governs the production of quantification. Rather, a market logic affects the way data are constructed, collected and compared. In this setting, measures are not merely assessed based on their epistemic qualities—for example, how well they capture the reality of higher education—but rather in their ‘market share’, i.e. the number of countries and agencies agreeing to participate and contribute to the work of measurement. In this chapter, we move away from the global level and examine the case of quality assurance in higher education in Europe in order to substantiate how, why and with effects this market of measurement works.

Finally, the book’s concluding Chapter 7 brings together the five different strands of the empirical and theoretical analysis, in order to argue for a novel perspective of the role of quantification in the production of education future utopias. The chapter discusses the ways that metrological realism has constructed a well-supported epistemic infrastructure, built on relationships and practices that go beyond the mere objectivity and reliability of numerical evidence; rather, quantification has become the new political imagination of planning and executing future education governing vistas.

REFERENCES

- Ball, S. (2007). *Education PLC: Understanding private sector participation in public sector provision*. Routledge.
- Bandola-Gill, J., Grek, S., & Tichenor, M. (2022). *Governing the Sustainable Development Goals: Quantification in global public policy*. Springer/Palgrave.
- Bauman, Z. (1992). *Intimations of postmodernity*. Routledge.
- Baumgartner, F. R., & Jones, B. D. (1993). *Agendas and instability in American politics*. University of Illinois Press.
- Beukel, E. (2001). Educational policy: Institutionalization and multi level governance. In S. Andersen & K. Eliassen (Eds.), *Making policy in Europe* (2nd ed.). Sage.
- Bhuta, N. (2012). Governmentalizing sovereignty: Indexes of state fragility and the calculability of political order. In B. Kingsbury, S. Merry, & K. Davis (Eds.), *Indicators as technologies of global governance*. Oxford University Press.
- Carayannis, E., & Campbell, D. (2011). *Mode 3 knowledge production in quadruple helix innovation systems: 21st century democracy, innovation and entrepreneurship*. Springer.
- Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. SAGE.
- Crowley, K., & Head, B. W. (2017). The enduring challenge of ‘wicked problems’: Revisiting Rittel and Webber. *Policy Sciences*, 50, 539–547. <https://doi.org/10.1007/s11077-017-9302-4>
- Desrosières, A. (1998). *The politics of large numbers*. Harvard University Press.
- Djelic, M.-L., & Sahlin-Andersson, K. (Eds.). (2006). *Transnational governance: Institutional dynamics of regulation*. Cambridge University Press.
- Etzkowitz, H., & Leydesdorff, L. (2000). The Dynamics of innovation: From national systems and ‘mode 2’ to a triple helix of university-industry-government relations. *Research Policy*, 29(2), 109–123.
- Eyal, G. (2019). *The crisis of expertise*. Wiley.
- 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.
- Funtowicz, S., & Ravetz, J. (1993). Science for the post-normal age. *Futures*, 25(7), 739–755.
- Gibbons, M., Limoges, C., Nowonty, H., & Schwartzman, S. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. Sage.
- Grek, S. (2010). International organisations and the shared construction of policy ‘problems’: Problematisation and change in education governance in Europe. *European Educational Research Journal*, 9(3), 396–406.

- Grek, S. (2008). From symbols to numbers: The changing governance of the European education space. *European Educational Research Journal*, 7(2), 208–218.
- Grek, S. (2022). The education Sustainable Development Goal and the generative power of failing metrics. *Policy and Society*, 41(4), 445–457. <https://doi.org/10.1093/polsoc/puac020>
- Grek, S., & Russell, I. (2023). Beyond Bologna? Infrastructuring quality in European higher education. *European Educational Research Journal*, 0(0). <https://doi.org/10.1177/14749041231170518>
- Hannerz, U. (1996). *Transnational connections*. Routledge.
- Hood, C., Rothstein, H., & Baldwin, R. (2001). *The government of risk*. Oxford University Press.
- Hooghe, L., & Marks, M. (2001). *Types of multi-level governance*. European Integration on Line Papers (EILP).
- Jasanoff, S. (Ed.). (2004). *States of knowledge: The co-production of science and the social order*. Routledge.
- Jessop, B. (2008). A cultural political economy of competitiveness and its implications for higher education. In B. Jessop, N. Fairclough, & R. Wodak (Eds.), *Education and the knowledge-based economy in Europe*. Sense Publications.
- Kenway, J., Bullen, E., & Robb, S. (2004). The knowledge economy, the techno-preneur and the problematic future of the university. *Policy Futures in Education*, 2(2), 330–345.
- Latour, B. (1987). *Science in action*. Harvard University Press.
- Lawn, M., & Lingard, B. (2002). Constructing a European policy space in educational governance: The role of transnational policy actors. *European Educational Research Journal* 1(2).
- Lawn, M. (2013). A systemless system: Designing the disarticulation of English state education. *European Educational Research Journal*, 12(2), 231–241.
- Logar, N. (2011). Scholarly science policy models and real policy, RSD for SCISIP in us mission agencies. *Policy Sciences*, 44(3), 249.
- Lövbrand, E. (2011). Co-producing European climate science and policy: A cautionary note on the making of useful knowledge. *Science and Public Policy*, 38(2011), 225–236. <https://doi.org/10.3152/030234211x12924093660516>
- Mayntz, R. (1994). *Modernization and the logic of inter-organizational networks*. Max Planck Institute.
- Milan, S., & Treré, E. (2019). Big Data from the South(s): Beyond Data Universalism. *Television & New Media*, 20(4), 319–335. <https://doi.org/10.1177/1527476419837739>
- Miller, P. (2001). Governing by numbers: Why calculative practices matter. *Social Research*, 68(2), 379–396.

- Miller, P., & Rose, N. (2008). *Governing the present: Administering economic, social and personal life*. Polity.
- Mulderrig, J. (2008). Using keywords analysis in CDA: Evolving discourses of the knowledge economy in education. In B. Jessop, N. Fairclough, & R. Wodak (Eds.), *Education and the knowledge-based economy in Europe*. Sense Publications.
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science: Knowledge and the public in an age of uncertainty*. Polity.
- Nowotny, H., Scott, P., & Gibbons, M. (2003). Mode 2 revisited: The production of knowledge. *Minerva*, 41, 179–194.
- Ozga, J., Seddon, T., & Popkewitz, T. S. (2006). *Education research and policy: Steering the knowledge-based economy*. Routledge.
- Pestre, D. (2003). Regimes of knowledge production in society: Towards a more political and social reading. *Minerva*, 41, 245–261.
- Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.
- Rafferty, J., Hanney, S., Greenhalgh, T., Glover, M., & Blatch-Jones, A. (2016). Models and applications for measuring the impact of health research: Update of a systematic review for the health technology assessment programme. *Health Technology Assessment*, 2016, 20.
- Ringel-Bickelmeier, C., & Ringel, M. (2010). Knowledge management in international organisations. *Journal of Knowledge Management*, 14(4), 524–539.
- Rosenau, J. (1999). Towards an ontology for global governance. In M. Hewson & T. Sinclair (Eds.), *Approaches to global governance theory*. State University of New York Press.
- Scott, J. C. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.
- Stehr, N. (1994). *Knowledge societies*. Sage.
- Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2009). Report by the commission on the measurement of economic performance and social progress. <http://www.stiglitz-sen-fitoussi.fr/en/index.htm>. Accessed on 22 January 2011.
- Thrift, N. (2005) *Knowing capitalism*. Sage.
- Tichenor, M. (2022). Statistical capacity development and the production of epistemic infrastructures. *Policy and Society*, 41(4), 541–554.
- 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*, 41(4), 431–444, <https://doi.org/10.1093/polsoc/puac015>
- Weick, K. E. (1995). *Sensemaking in Organisations*. Sage.

- Wesselink, A., & Hoppe, R. (2011). If post-normal science is the solution, what is the problem?: The politics of activist environmental science. *Science, Technology, & Human Values*, 36(3), 389–412. <https://doi.org/10.1177/0162243910385786>
- Whitty, G., Power, S., & Halpin, D. (1998). *Devolution and choice in education*. Open University Press.
- Yearley, S. (2005). *Making sense of science: Understanding the social study of science*. Sage.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Universality and Interdependence in Transnational Education Governance

I INTRODUCTION: INTERDEPENDENCE IN A COMPLEX WORLD

The dominance of International Organisations in the production of global metrics has not only penetrated the transnational social and policy fields; numbers have become an integral part of the fabric of International Organisations themselves. However, little is known about the ways in which global processes of quantification are reconfiguring not only the field of the global governance of education per se, but also—and crucially—the international organisations that have arguably brought it to existence. Metrics have infiltrated not only organisational cultures and the environments these organisations inhabit; crucially, they are reshaping the ways International Organisations co-exist, compete and survive in an increasingly quantified yet uncertain world.

Recent decades have seen fervent activity by International Organisations to build working collaborations and broad alliances for finding ‘global solutions’ to ‘global crises’. Financial investment in these collaborations is increasing and so is hope: *If only we had known, we could have acted*. Given the moral dimension that these new indices of progress have taken, as well as the enormous human and environmental cost of their failures, there is growing recognition that the interplay of International Organisations in transnational governance has led to the production of knowledge that is de-contextual, standardised, comparable and even at

times seen as *universal*, given its heavily moral undertones and the global nature of the discursive agendas it has given rise to.¹

Building on international relations (IR) theory, science and technology studies (STS), and using theoretical strands from organisational sociology, as well as the field of the social studies of metrics, this chapter examines the interrelationships of International Organisations (IOs) in constructing the global metrological field. As is well-known by now, IOs have been central to processes of standardisation, de-contextualisation and performance management through numbers; as a result, they have been instrumental in commensurating, and therefore transforming global education governance.

Thus, a central focus of this chapter is the—concomitant with the lure of numbers, albeit less spectacular—recent moves of large IOs not only to establish collaborative partnerships through connections with governments and local agencies, but also crucially with one another. The encoding of data processes and organisational cultures that these collaborative endeavours require (in order for data to be shared and co-produced) allows a comprehensive analysis of the workings of quantification for education governance. In other words, the examination of the interplay of IOs is a unique opportunity to open, rather than stack yet another ‘black box’ in the field of global monitoring (Bhuta, 2012).

Despite the renewed prominence given to the need for alliance-building by IOs, collaboration has always been central to their operation, since they have traditionally needed to work closely with governments, NGOs and the private sector. Yet, the complexity of ‘wicked’ problems, ‘donor duplication’, resource-pooling and data overload have become some of the most common reasons that IOs are increasingly compelled to work together. Indeed, most major global strategies, such as the Millennium Development Goals (2000–2015), the Sustainable Development Goals or major education testing regimes, such as the OECD Programme for International Student Assessment (PISA), are collaborative endeavours, dependent on pooling of resources and expertise. This chapter focuses on understanding the rise of universality of knowledge-making in transnational governance by focusing on the ways that IOs

¹ Here I follow Djelic and Sahlin-Andersson’s preference of the term ‘transnational’ versus ‘global’ governance, since ‘the label “transnational” suggests entanglement and blurred boundaries to a degree that the term “global” could not’ (2006, p. 4—for a more developed argument see also Hannerz, 1996).

learn from one another. In the making of numbers, how do they negotiate financial resources and knowledge controversies? How do they actively produce collective sense-making (Weick, 1995) and issue-framing strategies (Baumgartner & Jones, 1993)? How much do we know about their expert networks and their collaborations? Ultimately, if rating and ranking practices are a ‘zero-sum’ game for the assessed, how much do we know about the rules of the game for the assessors?

Empirically, the chapter examines two separate cases from the field of education. Education policy, both in the global South and the global North, has increasingly been dependent on the measurement of its performance for the improvement of human capital. Education can be a productive vantage point, since assessment and quantification of performance have a very long history in the field. It is a key element in the newly emergent well-being and sustainability strategies that have prevailed the statistical governing project post the global pandemic. Education is closely congruent with the efforts to use ‘softer’ data sets for calculating the social. Last but not least, it is one of those policy areas that large IOs like UNESCO, the OECD, the European Commission and the World Bank have invested large amounts of data and expertise from the mid-twentieth century.

The chapter begins with a short review of the literature of the politics of quantification; it then moves on to a consideration of the theoretical underpinnings of the analysis and continues with the presentation of the two case studies under examination. Finally, it is concluded by a discussion of international organisations, interdependency and metrology in the field of transnational education governance and beyond.

2 ‘GOVERNING BY NUMBERS’ IN TRANSNATIONAL GOVERNANCE

Scholarship on the role of numbers in governing societies has been abundant and has attracted multiple fields of study, including sociology, history, political science, geography, anthropology, philosophy, STS and others. Prominent authors have written lucidly about the role of numbers in the making of modern states and the governing role of measurement regimes in various areas of public policy and social life (Alonso & Starr, 1987; Desrosieres, 1998; Espeland & Stevens, 2008; Hacking, 1990, 2007; Porter, 1995; Power, 1997; Rose, 1999). Similarly, anthropologies of numbers suggest 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, p. 23; see also influential work by Merry, 2011; Sauder & Espeland, 2009; Strathern, 2000). Further, important insights and perspectives on indicators in particular come from STS (Bowker & Star, 1999; Lampland & Starr, 2009; Latour, 1987; Saetnan et al., 2011), including actor network theory (Latour, 2005). Finally, there is a small but growing body of studies relating to specific uses of indicators and quantification in transnational governance contexts (for example, Bogdandy et al., 2008; Palan, 2006; Martens, 2007; Fougner, 2008; Bhuta, 2012).

Nonetheless, despite the burgeoning number of publications on the global ‘governing by numbers’, our understanding of the relationship of the politics of measurement and the making of transnational governance is less well-examined; as Djelic and Sahlin-Andersson (2006) suggest, due to the fluidity and complexity of the intense cross-boundary networks and soft regulation regimes that dominate the transnational space, transnational governance is a particularly productive field of enquiry on the role of numbers in governing. This lack of attention could be due to disciplinary boundaries; for example, scholars of IR and international law have not paid much attention to the field so far, although there is a rise in some interesting literature of the role of numbers in global political economy (for example, Fougner, 2008; Martens, 2007; Palan, 2006).

What are the properties of numbers that would suggest such a central role in the production of transnational governance? By contrasting numbers to language, Hansen and Porter (2012) 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. By using the example of the barcode, they lucidly illustrate ‘how numerical operations at different levels powerfully contribute to the ordering of the transnational activities of states, businesses and people’ (2012, p. 410). They suggest the need to focus not only on the nominal qualities of the numbers themselves but also, according to Hacking, ‘the people classified, the experts who classify, study and help them, the institutions within which the experts and their subjects interact, and through which authorities control’ (2007, p. 295).

It is precisely on international organisations as data experts that this chapter focuses upon; following the literature on the capacities of numbers to both be stable yet travel fast and without borders, the chapter sheds light on what Latour called ‘the few obligatory passage points’ (1987, p. 245): in their movement, data go through successive reductions of complexity until they reach simplified enough state that can travel back ‘from the field to the laboratory, from a distant land to the map-maker’s table’ (Hansen & Porter, 2012, p. 412).

3 THEORETICAL FRAME AND KEY INTERMEDIARY CONCEPTS

The chapter follows a ‘constructivist-institutionalist’ approach (Smith, 2009), as it works with Lagroye’s definition of governing as ‘a set of practices which participate in the organization and the orientation of social life’ (1997, p. 25). Thus, it builds on the premise that far from being a system composed uniquely of ‘national’ and ‘transnational’ bodies, *governing the transnational* is an ‘institutional order’ made up of all the actors who participate in the construction and institutionalisation of global problems (Smith, 2009). In turn, transnational ‘governing’ is conceptualised as those ‘assemblages of apparatuses, processes and practices’ that make governing happen (Clarke & Ozga, 2011).

As already suggested, a considerable body of research has already focused on the work of IOs in transnational governance. Yet, this research has often seen them as monolithic institutions, or actors with similar interests in a similar context, without attention to the complex set of realities that bring them together and apart over time (with notable exceptions of course, see for example Cram, 2011). IOs are often also seen as *internally* stable—this means that divisions of authority, institutionalised norms, expectations and values are thought to be commonly shared by all actors within an IO. Nevertheless, ‘most of the time, [...] at least some of the actors within an IO will be seeking to change at least some of its institutions, whilst others will work to retain their stasis’ (Jullien & Smith, 2010, p. 4). The examination of actor alliance formation and mobilisation is hence vital in order to understand these relations—both upstream, i.e. the setting of rules and problem framing, as well as downstream, namely the application and maintenance of rules among the actors who are all engaged in competitive relationships (Jullien & Smith, 2010).

Indeed, some of this actor mobilisation and alliance-building is achieved not internally but through networking with other IOs.

Thus, one of the key concepts that mobilises this argument is the notion of ‘political work’ (Smith, 2009). When one studies political work, institutions themselves are not the objects of study per se; rather, the focus of the investigation is on the continual cycle of institutionalisation, deinstitutionalisation and reinstitutionalisation of ideas and values within the organisation in question. The study of quantification as a policy instrument can become a particularly fruitful context for such an analysis as one can examine ‘political work’ as those processes that engender the construction of new arguments and the activation of new alliances; subsequently, they either produce change or reproduce institutions, namely actors’ rules, norms and expectations (Jullien & Smith, 2010).

Before moving on, our attention needs to be directed to two intermediary concepts, those of the ‘field’ and ‘knowledge controversies’. To start with the latter, Barry (2012) uses the notion of ‘political situation’ to explain the ways that STS could have been misguided in their definition of knowledge controversies as conflicts that relate principally to a clash of scientific evidence and ideas. Instead, he suggests that ‘the significance of a controversy needs to be understood in relation to a shifting and contested field of other controversies and events that have occurred elsewhere and at other times’ (Barry, 2012, p. 324). Whereas STS initially mostly focused upon the ‘black box’ of science by looking at issues of credibility, objectivity and reliability (Shapin & Schaffer, 1985), it then moved on to the analysis of public knowledge controversies, where expert knowledge would clash with public, lay knowledge (Wynne, 2003). Yet, Barry argues that despite the growth of transnational standardisation processes, the issue of knowledge controversies has not been addressed either by the IR or the STS literature, as if the simplification of data (and the consensual expert practices it involves) decreases rather than increases the possibility of knowledge disputes and failures. However, it is widely known that achieving transnational standards is infinitely difficult; political contestation often gets submerged and hidden behind the need to appear as collaborative and open to partnerships. Knowledge contestations are then seen as an impediment to the formation of collaborations. In fact, it appears that it is precisely in the knowledge controversies that one has to focus upon, if one aims to understand the very process of simplification and the exclusion of unwieldy or awkward data (or awkward experts for that matter). To return to Barry then, ‘what the concept of political

situation captures, is how the significance of a controversy is not so much determined by its specific focus, but needs to be conceived in terms of its relations to a moving field of other controversies, conflicts and events, including those that have occurred in the past and that might occur in the future' (2012, p. 330).

Second, the chapter suggests the need to examine the interplay of IOs as they construct the 'global metrological field'. Emanating from physics, the notion of field has been used in the social sciences in order to broadly refer to actors' relational topographies. Nevertheless, it is often reduced to merely looking at specific geographical and relational spaces. Yet, as Djelic and Sahlin-Andersson also suggest (2006), such a conceptualisation of fields misses a vital ingredient from the way fields operate; that is an understanding of the field as a field of power. Drawing on Bourdieu, transnational governance appears as a field of actors who constantly negotiate and push their own agendas forward; according to Bourdieu (1993), the logic of positionality is what gives the notion of the field meaning. In other words, the positions occupied by the different agents in the field, their advances and withdrawals relate to their efforts for distinction within this field as an expression of their professional, educational, or other interests. Meanwhile, the structure of the field is neither static, nor does it change in any systematic way. On the contrary, it is endlessly reformulated, according to the agents' struggles for recognition and improvement of their situation. Agents use the force of their economic, social, cultural or epistemic capital to raise their game and advance their front. It is the *relational* nature of these advances that gives the field its explanatory significance. Thus, following Bourdieu, the chapter uses Djelic and Sahlin-Andersson's idea of fields as 'complex combinations of spatial and relational topographies with powerful structuring forces in the form of cultural frames or patterns of meaning' (2006, p. 27). An examination of the interplay of IOs in the rise of the global metrological field is therefore necessary, as it is vital to examine transnational governance not only as a field of numbers or as a field of actors, but also as both.

Thus and to conclude, the chapter adopts a constructivist standpoint by focusing on the social and political conditions that influence the production of numbers, adopting the ontological position that their existence is not organic but rather the product of the interconnectedness of IOs, as outlined above. It examines IOs for whom this transnational game exists ('what keeps them running' as Bourdieu would put it) and even

national actors who just utilise it as an instrument in their local political battles.

4 THE CASE OF THE OECD AS A EUROPEANISING ACTOR: THE RISE OF THE INTERDEPENDENCY OF THE OECD WITH THE EUROPEAN COMMISSION

How do international education agendas look like from the perspective of national education systems? The answer is that they look broadly similar, without much attention being paid to the source of expertise and policy recommendations, as long as they fit with the local policy agendas and direction of reforms. Indeed, the empirical analysis in this section broadly builds on previous research (Grek, 2009, 2012, 2020) that suggested that the European Commission (EC) and OECD recommendations are often received at the national level as homogeneous. In order to understand this increasing trend, the empirical investigation moved beyond top-down accounts of the mere and one-directional transfer of policy from the international to the national, towards more attention to the interaction and mediation across 'levels' and actors. In terms of methods, the empirical analysis focuses on the examination of policy discourse as well as the interviews of 15 actors from both the Commission and the OECD, as well as other relevant research agencies; the interviews focused on the actors' role in processes of coordination (conferences, meetings, project work), their interactions with other actors within and beyond their organisations and other relational ties that link them and others through channels of flow of data, ideas and/or material resources. The policy actors interviewed have had positions of power and significant decision-making leverage: they had first-hand experience and participation in meetings and debates between the Directorate General Education and Culture (DGEAC) and the OECD in regard to the financing and conduct of large international assessments.

Hence, this chapter examines how the OECD became a dominant education policy actor as a result of its deliberate and systematic mobilisation by the European Commission which found in the OECD not only a great resource of data to govern (which it did not have before) but also a player who would be pushing the Commission's own policy agenda forward, albeit leaving the old subsidiarity rule intact. As I will show, testing is important here because it produces numbers and consequently

ratings and rankings; once the OECD had created PISA's (Programme of International Student Assessment) unprecedented spectacle of comparison in European education, no system could remain hidden any longer. The field of measurement became instantly the field of the game.

4.1 *International Comparative Assessments: the OECD's International Adult Literacy Survey (IALS) and PISA*

Indeed, international comparative testing has become the lifeblood of education governance in Europe and globally. It is more than simply a project of measurement; rather, it has become part of consistent efforts to restore legitimacy and trust between populations and their governments. As Hall contends, 'building legitimacy requires potential users in the process, as well as technical experts. The most important role of indicator sets may be in framing the issues and defining the problems, rather than suggesting the solutions' (2009).

The governance of international comparative testing reflects these values. Project boards usually work in conjunction with a large range of consortia of international partners and technical advisors (statisticians, media specialists and, interestingly, philanthropists); they also consult with a vast array of different actor groupings, such as academics, private companies, policymakers, associates, country correspondents, regional working groups and others. Regular training courses are delivered as well as seminars, and regional, thematic and global conferences. Although all these initiatives suggest sustained efforts to include and create consensus with the greatest number of stakeholders possible, the role of experts remains central: before they acquire a more 'public' and visible face, tests are being discussed, negotiated and indeed fought over among field experts for a long period of time.

The case of the OECD as a knowledge producer for education governance is particularly interesting because, unlike the EU, it has neither the legal instruments nor the financial levers to actively promote policy-making at the national level within member nations. Nonetheless, through ranking exercises such as the 'Education at a Glance' annual reports, the Indicators in Education project (IALS), through PISA and its national and thematic policy reviews, its educational agenda has become significant in framing policy options not only at the national but also, as it has been argued, in the constitution of a global policy space in education (Grek & Lingard, 2023; Lingard et al., 2005; Ozga & Lingard, 2007).

This raises the question—what transformed the OECD into one of the most powerful agents of transnational education governance? What are the qualities of the OECD’s expert work that maintain the organisation as a highly trusted source of education policy recommendations both prior to and post-COVID? Martens (2007) has contributed substantially to this discussion suggesting that the ‘comparative turn’—‘a scientific approach to political decision making’ (2007, p. 42)—has been the main driver of OECD success. Through its statistics, reports and studies, it has achieved a brand which most regard indisputable; OECD’s policy recommendations are accepted as valid by politicians and scholars alike, ‘without the author seeing any need beyond the label “OECD” to justify the authoritative character of the knowledge contained therein’ (Porter & Webb, 2004).

Drawing on Marten’s (2007) ideas, we can see that by now there is a taken for grantedness about education indicators, despite all the commentary asking for contextualisation in their interpretation (e.g. Nóvoa & Yariv-Mashal, 2003), and this is indicative of the way in which they have become an accepted part of the contemporary educational policy lexicon across the globe, within and well beyond the OECD, and of their growing significance to the work of the OECD itself since the 1980s. Despite its ups and downs, and the supposed demise of its glamour, PISA continues to account for a large chunk of the Education Directorate’s budget inside the OECD. One could suggest that the OECD’s greatest impact has been in relation to its Indicators agenda, including PISA, and its role in constructing a global educational policy field through governance by comparison (Martens, 2007; Ozga & Lingard, 2007). Indeed, Antonio Nóvoa argued, ‘comparing must not be seen as a method, but as a policy ... the expert discourse builds its proposals through “comparative” strategies that tend to impose “naturally” similar answers in the different national settings’ (2002, p. 144). Although that might be too stark a claim, and although comparison can be both (there are certainly good epistemological reasons for comparative research that owe nothing to policy), it is still important to acknowledge the sustained power of comparison as a governing technology, especially when governing is done at a distance and through the use of ‘soft power’.

Thus, a brief historical analysis would show that there has been a range of such studies that the OECD has been organising since the early 1990s, the majority of which were adult literacy studies to start with, followed by the delivery of the most successful one, PISA and PIAAC, the Programme

for the International Assessment of Adult Competencies. The first International Adult Literacy Study (IALS) was the first and largest international comparative testing regime of its kind. Conducted from the early 1990s, IALS was innovative, as it was the first time ever that an international comparative dimension was added to the construction of a literacy survey instrument. Thus, it heralded a new era in the construction and evolution of international comparative studies, as for the first time ever it gave international testing a comparative dimension, where measurement against other countries' performance offered unprecedented visibility and thus exposure. As it was an original and new endeavour, slowly at the start but increasingly later on, IALS boosted confidence in the construction of measurement tools of this kind, increased their persuasive power in regard to their validity and transparency and created substantial revenues to the research agencies administering them. Finally, and perhaps above all, it created a circle of like-minded expert communities, who found in these studies a platform for promoting the problematisation of specific issues, their institutionalisation through their exchanges and the setting up of the study, as well as their legitimisation, in the form of advice to failing countries, once the results were published.

Following the successful IALS endeavour, the Programme for International Student Assessment (PISA) became a major instrument in providing data for the European education systems almost from the start. The international dimension of the survey, which overrides the boundaries of Europe to compare student performance in countries as diverse as the United States, Greece and Indonesia, gave PISA a particularly significant weight as an indicator of the success or failure of education policy. While always testing reading, mathematical and scientific literacy, its innovative dimension—and part of its interest as a governing device—lies in the fact that it does not examine students' mastery of school curricula, rather the focus is on an assessment of young people's ability to practically apply their skills in everyday life situations. The focus on 'real-life' circumstances and on students' capacity to enter the labour market with core skills, such as literacy and numeracy, has taken PISA's focus of interest away from less explicit educational aims that resist measurement (e.g. democratic participation, artistic talents, understanding of politics, history, etc.), towards a more pragmatic view of education's worth: 'its relevance to lifelong learning' (OECD, 2003). Finally and perhaps most significantly, a key feature of PISA is:

its policy orientation, with design and reporting methods determined by the need of governments to draw policy lessons. (OECD, 2003, no page numbers)

Hence, this is not simply a testing regime—it is constructed and operates under a clear and specific policy framework, which is to be adopted by the participant countries if they are to improve their future PISA assessments and thus improve their standing in attracting economic and human capital investment. In other words, the involvement of the OECD with the steering of education policy in participant countries does not stop with the publication of the PISA—or whichever study’s—results; on the contrary, this is perhaps where it begins. Expert groups write expert reports, analysed and taken forward by other national and local experts, while the Commission expert committees are also on board in order to keep the game in sight and keep it running.

This is the kind of status that the OECD acquired with the conduct of large international tests; the seal of unequivocal and trusted expert truth. In other words, OECD not only produces evidence quickly and effectively but also digests it and offers it to policymakers in the format of policy solutions. In a sense, if we are used to accounts of European policymaking as slow, cumbersome and ‘coming from nowhere’ (Richardson, 2001, p. 21), the OECD bypasses these obstacles in four key ways; first, it defines the limits of the possible by suggesting what can be measured, hence what can be ‘done’; second, it carries no political jurisdiction, therefore, it carries no external threats to national policymaking, as perhaps the Commission or other EU institutions might have done; it now has the experience, networks and the technical and material resources to speed up the policy process so that it can show ‘results’ within the usually short timeframe that policymakers are in power; and last but not the least, it carries all the ‘right’ ideological messages for education systems in the twenty-first century—that is, it connects learning directly to labour market outcomes and economic growth.

Nonetheless, despite the numerous and in-depth analyses of the OECD’s education measurement work, how can the OECD continue to be such a powerful player in education governance in Europe? As some of the people who work there might have argued, the OECD Education Directorate staff who are based in Paris take a few decisions, if any; the OECD, as they argue, is no other than the participant countries and

the national actors and experts sent to the OECD committees and meetings. Thus, how accurate is to examine the emergence of this new policy arena by simply focusing on this single international actor? Not entirely accurate, as the following interviewee suggests:

So around 2003–04, we [OECD and Commission] started becoming far more involved. Meetings all over the world, I don't know how many countries I visited but what is important is that the Commission is there.... The European member states should see that the Commission is there because one of the criticisms of the Commission since all this started was that we didn't take into account all the good work of the OECD. Which was wrong but they said it. The way of showing them was to actually be there – not an empty chair. (EC4)

Indeed, although the Commission and the OECD had been leading quite separate ideological paths, a new 'love affair' between the two organisations began emerging—this relationship would gradually strengthen and eventually become the *sine qua non* for the governing of European education systems. Another interviewee was even more eloquent in his discussion of this flourishing relationship:

We used to have great competition between the two institutions [OECD and the EC] which was that they were research-based, we were policy-based. And we needed that. They needed the policy aspect to mobilise the European consciousness ... it was in their interest working with us ...We had some differences but we are working closer and closer together, we are very very good friends now, there is no conflict. (EU3)

As evident in the following quote, the relationship soon became more than one of influence:

When the OECD started speaking about TALIS [survey on teachers] it attracted the attention of the member states that all this is very good but it is expensive. ... So I managed to convince my Director General of supporting (the OECD) with an awful lot of millions of euros. And I went back to the OECD with that message and said that of course if we pay we want influence. (EC7)

On the other hand, OECD actors appear also as quite open to the Commission, stressing from their own point of view, the reasons that the DG Education would work closely with them:

First of all I think we've been very lucky that on the Commission side, that they've given a lot of emphasis to skills recently and they have this 'New skills for new jobs' initiative and so I think we were fortunate that the work that we decided to do on PIAAC corresponded extremely well with their areas of interest and research priorities... So they made a direct contribution, an actual contribution to the international costs and also eventually agreed to subsidise EU countries, the cost that they had to pay as well to the OECD. So we got just a block of direct funding and indirect funding to countries that they then had to pay us for the international costs. That made a big contribution in financial terms and therefore of course enhanced interest in the project. (OECD3)

Another OECD actor also suggested the way that the relationship, rather than hostile, has been much closer recently, in fact 'hand in hand':

We have the same perceptions like other international organisations that it is important that we work together and that we avoid duplication of effort and that we know what the other organisations are doing and that there are often occasions that jointly we can do more than what we can do individually. I think we were always aware of that but I think that has become increasingly important that we work hand in hand. And inevitably because we have some common goals. The OECD has had for some time its own job strategy, the Commission has its own employment strategy and its Lisbon goals and there is a lot of overlap. So, I think it is quite normal that we can cooperate on a lot of areas. (OECD5)

However, there is also a reverse side to the coin. Even though the interdependency and collaboration between the two organisations increased, more often than not these exchanges take place in a competitive field, where the delivery of studies and the collection of education statistics is not a choice anymore, but a necessity. Conflict and tensions can run deep:

The main reason is that they are competitors and both in scientific and in financial terms it is getting more and more difficult to conduct these surveys. There was a message from member states to the OECD and the IEA – get together, sit down and discuss it and do it. Now, 6 months

later, we all come together and we ask what was the result of that meeting and the answer was that we didn't find a date. They don't work together because they don't like each other. (EC9)

Interviewees also describe internal conflict within international organisations and their departments, for example, within the OECD itself. The following quotation describes the conflict between the Centre for Educational Research and Innovation (CERI) and the Directorate of Education, similar to the kinds of processes Jullien and Smith (2010) describe when they discuss IOs as internally unstable institutions, rather than the opposite:

They live in different worlds – the same floor at the OECD but in different worlds. They don't like each other – one is more research-based, the other one more indicators and data, surveys. One is more reflection, the other one is more publicity, the charts – different traditions, the same director. (EC12)

Finally, another account that describes the conflict and competition for securing contracts for education research in Europe comes from another interviewee, a key member of staff of one of the Commission's research agencies:

I think because the OECD is very much looking for member states' subsidies and grants and financial support for each separate research activity, they are also keen in showing that they do something unique and innovative in order to get such funding. And so then in a way they are in competition with us. An example is they did a recent policy review which is called 'Learning for Jobs' which basically deals with VET. And they didn't invite us to some national expert groups and so on that are in development—and they did very little use of our work because they wanted to do something that was different and specific so that they could sell it to the member states – this is my interpretation, of course. But I think that there is this kind of competition, differentiation between European institutions because we are in competition for funding. (EC3)

The quotations presented above suggest that the descriptions of a field of actors who come together harmoniously and in partnership to work on certain agendas might be misleading. On the contrary, they highlight the need to also focus our attention on those meetings that never

happen, as well as those actors who are consistently not invited to expert meetings. They direct us to an understanding of a field, which is riddled with internal and external competition for funding, especially in times of reducing national budgets in an era of crisis. What is important to note is that the story of the collaboration of the OECD with the European Commission is neither a story of smooth consensus and collaboration nor merely a story of struggle and competition: rather, it is the unfolding story of continuous and increasing *interdependence of the two organisations in the production of governing knowledge*. The latter is both strategic and technical and functions as a way of creating the conditions for a *universal and consensual way of amalgamating data and policy priorities*, as the next empirical case will discuss.

5 FROM DISSENSUS TO UNIVERSALITY: THE CASE OF NARRATIVE-MAKING IN THE SDG4

The previous section focused on an analysis of the ways quantification increased the interdependence of international organisations through the construction of a single measurement field; we analysed the specific political work that PISA and other international assessments achieved by bringing together actors around common ideas and measurements, even in instances of conflict and disagreement. This section will move this discussion further, in order to show how this increasing interdependence in the construction of a single global metrological field in education has created the conditions for the emergence of common, universal ideas (and ideals) of what education is for and what it should achieve. For this analysis, we will turn our lens to the production of the Sustainable Goal 4, the education SDG that has proclaimed to ‘ensure inclusive and equitable quality education and promote lifelong learning opportunities for all’ (UN, 2023). Specifically, we will focus on a discourse analysis of two crucial documents that emerged in the preparatory stages of putting together the SDG4: these are the 2014 Muscat Agreement and the 2015 Incheon Declaration.

Using discourse analysis of these documents, I will discuss 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 does not only bring actors a lot closer together than ever before (as discussed in the previous section), but is also inscribed, materialised and made plausible by the production

of a powerful—however ambiguous—rhetoric of development, equality, democracy, universality and morality. If statistical data is all about the measurement of possibility, the construction of universal narratives of progress fosters *plausibility*; narratives bring coherence and give meaning to informal and fragmented global governing spaces. In the following sections, I will 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 re-frames them as global goal-setting endeavours. The aim of this analysis is also partly conceptual; the examination will show how the emergence of the SDG4 narrative requires the 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.

5.1 *The 2014 Muscat Agreement*

As indicated in the Introduction, the global governance of education from the start of the twenty-first century was characterised by the coexistence of multiple, and sometimes overlapping arguments about the role of education in society: these arguments were not always harmonious or conflict-free. On the contrary, there have been significant power asymmetries and competing expectations in relation to aims and objectives of the policy priorities in discussion, as well as the decision-making architecture towards their realisation.

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. 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) as a result of 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 UNESCO-led EFA negotiations faced competition by 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) was very much the

product of consensus, carefully crafted by the global education community and reflected the multiple priorities of education agencies while also allowing civil society to make a meaningful contribution, this was not the case with regard to the MDGs, which viewed education in much more narrow terms and focused exclusively on universal primary education.

Since the negotiation of the SDGs was approached by different agents as an opportunity to put an end to this duality of education agendas (given that disagreement might have meant an exclusion of education from the SDGs), the re-alignment of the EFA agenda with the MDG education ‘camp’ required crafting a new set of education targets. Thus, it was through a new agreement on the goals that a break-through was found: this was the Muscat Agreement, signed in May 2014. The document was approved 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—i.e. 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, was 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 also is 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. (GEM, 2014, p. 1)

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 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, as according to the document, 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 Millennium 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. (GEM, 2014, p. 1)

The use of evidence in narrative-making is a powerful rhetorical tool in creating the necessary epistemic and measurement context 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 future of universality 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. (GEM, 2014, p. 2)

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; 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' (GEM, 2014, p. 2, my 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 in the last; second, the challenges of disagreement and of creating some form of consensus in a really complex and conflicting field; 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.

However, where is it that the education community should now move to? What kind of governing instrument can ensure unity and create a universal agenda, accepted by all participating actors? As the document continues, goal-setting becomes a 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 a 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 nominal targets without numerically specifying them:

We support “Ensure equitable and inclusive quality education and life-long 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.... (GEM, 2014, p. 3)

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 a numerical narrative, even without the use of specific numbers—just notional percentages (to be agreed) of an imagined world ‘by 2030’. This is how numerical narratives construct universal agendas of education progress, 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’ (GEM, 2014, p. 3). 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 script becomes the governing locus where organisational, epistemic and political struggles manage to settle. Its significance is evident, by the documents that succeeded it; first of which was the Incheon Declaration.

5.2 *The 2015 Incheon Declaration*

The Education 2030 Incheon Declaration was published in the World Education Forum, in Incheon, the Republic of Korea, from 19–22 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’ (UNESCO, 2016, p. 5).

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 Millennium 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. (UNESCO, 2016, p. 22, 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 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 to 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’ (UNESCO, 2016, p. 24). 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 also 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. (UNESCO, 2016, p. 32)

This—importantly—is not only a narrative outlining the policy direction, but rather it also offers 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 (the Muscat Agreement, for example) 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’ (ibid., p. 35).

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, but 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’. (UNESCO, 2016, p. 36)

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 (versus 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, but 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, p. 9)

To conclude, it is evident that the SDG4 is not exclusively a performance monitoring agenda. It uses a strong narrative built around it, in relation not only to the need for measurement towards achieving the priorities set (described almost exclusively in the language of different sets of indicators), but also for the new agenda to be seen as necessary, ethical, participatory and universal.

6 INTERNATIONAL ORGANISATIONS: INTERPLAY AND INTERDEPENDENCE IN THE MAKING OF THE GLOBAL METROLOGICAL FIELD

Through the analysis of the collaboration between the OECD and the European Commission, as well as the co-production of education indicators for the SDG4, this chapter has evidently shown how IOs do not constitute ‘centres of calculation’ independently from one another. Increasingly we find that they need to collaborate in the production of global education metrics. However, according to Merry (2011), IOs are not significant merely in terms of their knowledge production capacities, be they combined or separate. By examining specifically the role of indicators in transnational governance, Merry elucidates the governing effects of numbers themselves. Consequently, if we consider IOs central in the production of knowledge, we can infer that their operation—as the knowledge gatherers, controllers and distributors—must have crucial governing impact (2011). These effects empower IOs and set them in a complex and ever-evolving power game for influence and resources. Through an examination of the interplay and interconnectedness of IOs’ data apparatuses, it is precisely this power game and its rules that this chapter tries to cast light upon. Indeed, Shore and Wright argue that, ‘while numbers and “facts” have both knowledge effects and governance effects, it is also important to consider how these are produced, who designs them, what underlying assumptions about society shape the choice of what to measure, how they deal with missing data, and what interests they serve’ (2015, p. 433).

In light of this chapter’s empirical analysis and in the tradition of the seminal work of Barnett and Finnemore (1999), we need to question the International Relations’ conceptualisation of IOs as passive entities which merely distribute ‘principles, norms, rules, and decision-making procedures’, as the more economic, rational-theory analysis would have seen them to be. Instead, building on sociological institutionalism, Barnett and Finnemore see IOs as powerful agents which have ‘power independent of the states that created them’. Thus, they are purposive actors (Cox, 1992, 1996; Haas, 1992): ‘they define shared international tasks (like “development”), create and define new categories (like “refugee”), create new interests (like “promoting human right”), and transfer models of political organization around the world (like markets and democracy)’ (Barnett & Finnemore, 1999, p. 699).

However, given the prominence of IOs in IR literature, it is surprising how little attention has been given to the interplay, organisational overlaps and mutual dependencies of IOs. As this chapter has shown, rather than state-bound, IOs are increasingly dependent on other IOs to operate. For example, we find that new IOs are usually founded by other IOs, rather than member states (Shanks et al., 1996). In addition, staff mobility in IOs is very high: ‘a large part of staff ...is employed on fixed term contracts which generally run up to three years with the possibility, but not the obligation, of renewal’ (Ringel-Bickelmeier & Ringel, 2010, p. 525). In fact, the case of the OECD is particularly interesting, since it has ‘annual turnover rates sometimes as high as 40 per cent for certain staff’ (Ringel-Bickelmeier & Ringel, 2010, p. 526). The ‘revolving doors’ of IOs may suggest that staff often move between them, or even occupy multiple positions at the same time.

Hence, organisational interplay matters. Although, as Brosig (2011) suggests, IOs are dependent on states, the case of the rise of the global education policy field shows clearly that IOs do perform operations that states cannot and will not perform—in fact, most of them were founded in order to operate as cross-governmental diffusers of knowledge and norms. Barnett and Finnemore are again helpful in suggesting that cooperation between IOs may create mutual dependency, a situation that IOs would normally be seen to want to avoid (1999). Nonetheless, given the complexity of transnational governance and the technological advances of the last decade, we are facing a different situation altogether: IOs cannot and do not act independently to solve major social problems and challenges. Hence, and as the case of the education policy arena has shown, we see IOs as increasingly mobilising 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 complete insularity and autonomy (Raustiala & Victor, 2004). In addition, as we saw above, IOs are characterised by highly mobile workforces; what does this increased actor density and fluidity suggest about the coordination of measurement practices? Indeed, it appears that states ask for the collaboration of IOs as it is seen as a way of increasing efficiency, resource-pooling and coordination of their agendas—the example of the ways that European Commission’s DG Education and Culture was in effect compelled to work with the OECD

because of efficiency concerns by the member states, is a good one here (Grek, 2009, 2014).

The concept of organisational interplay is not entirely new to IR: there has been some stimulating work that has examined the interplay of ‘international regimes’ and consequent attempts to produce typologies (Gehring & Oberthür, 2009; Raustiala & Victor, 2004). Nonetheless, regimes lack precisely what Barnett and Finnemore (1999) suggest above: agency (Rittberger & Zangl, 2006). However, even when IR theory has acknowledged IOs constitutive nature as actors, there are other problems. By examining treaty regimes, for example, Young suggested two typologies for organisational interactions: nested and overlapping institutions (1996). But, as Brosig (2011) suggests, ‘research on regime complexes in which relations between institutions are of such density has indicated that disentangling them would compromise the collective character these regimes have acquired’. In addition, most of IR theory that has examined treaty regimes has done so from a rational-theory perspective, one that would explain the interactions as serving specific IOs interests and benefit calculations (Galaskiewicz, 1985; Oliver, 1990; Van de Ven, 1976).

Nonetheless, even when IOs are assigned with agency, asymmetries and power relations are only explained on the basis of rational, interest-based behaviour. However, as the example of the European Commission’s collaboration with the OECD has shown (Grek, 2009, 2014), material resources do not always explain organisational interaction; IOs may actually be very well-off but lack the knowledge and expertise, even legitimacy to promote specific policy agendas. To use the same example again, the notion of subsidiarity would suggest that for the European Commission, the OECD could act as a mediator of its own policies in the member states. In other words, DG Education and Culture lacked the legitimacy to enter national policy spaces; OECD, as an expert institution, did not. On the other hand, organisations like the OECD may well have both the resources and the expertise, but could be lacking in policy direction and influence.

Therefore, it is evident that although important scholarship in the fields of IR, organisational sociology and the social studies of quantification exists, little has it enlightened us about the politics, processes and practices of the interdependence and interplay of IOs in the field of the production of global metrics. On the one hand, IR theory has emphasised the role of IOs in transnational governance; initially through an examination of treaty regimes, and later with an emphasis on IOs influence in

power play, the field is dominated by rational, interest-based theoretical perspectives. Thus, it has failed to examine qualities of IOs that relate to their constitutive powers as independent, yet interconnected, actors in the shaping of global policy agendas through their expert knowledge work. On the other hand, organisational sociology, although rich in its intellectual history of competing views about how organisations work, has not as yet examined closely the role of numbers in reshaping organisational behaviour. The insistence on separating the internal from the external organisational ‘lifeworlds’ fails to take into account precisely what this chapter has shown that numbers are able to do: that is, diffuse boundaries and set IOs in a more complex and fluid reality. Finally, studies of quantification, although coming from a wide variety of disciplinary perspectives, have largely focused on the role of numbers as agents in themselves; there has been little, if any, attention, to the political work of the actors that organise these processes, that shape and are shaped by them. The chapter has shown how a focus on the interdependency of international organisations in the production of education metrics and narratives is a useful notion for explaining the slow build-up of a global governing infrastructure, made up by a variety of actors and using education data as its connective tissue.

However, if this growing interdependence of IOs on transnational education governance has significant effects, what do these effects look like? How can we trace the ways that these new collaborations and exchanges change the global education policy field? This chapter examined two major collective declarations as the production of a universal narrative of progress that brings all relevant actors together as participants of a single technocratic exercise. As the discourse analysis of these policy documents shows clearly, although the myth-making of education utopias where ‘no-one is left behind’ has always been prevalent in education, what quantification has managed to deliver is the transformation of vague and broad premises into a technical programme of prescribed numbers and targets, crafted and co-signed by a range of actors. In other words, it enveloped former utopian political aims with the material inscriptions of numbers, modelling therefore not only what such education futures should look like, but also *how* to achieve them. Thus, the chapter showed how numerical narratives are used as the material building blocks of a governing infrastructure where all major IOs and other types of actors

converge. Thus, numerical narratives become 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, p. 770).

Indeed, it is the potential of such numerical 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 education 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’ (1984, p. 10). The intelligibility of events, actors and decisions is of particular significance in global education governance, 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 also 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.

Additionally, I have also discussed the ways that stories and narratives in global education 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 universality: they are not only the valuable resource that allows actors to understand—even feel—the emergency, but also through the dominant instrument of goal-setting, metrics are also offered as the vital component of establishing universal ideals of education futures. Here, I 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, in constructing causality by offering 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 impose coherence on complex and messy political realities—and they do so predominantly by selecting ideas and events

that are organised by the chosen plot while 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, as well as 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 also 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.

To conclude, this chapter focused on the materialities and intertwinements of numerical data, actors and discourse. In doing so, it manifested their vital work in creating actor interdependence and ultimately the universality of datafied education governance, as the only viable means of achieving consensus and a technical/political equilibrium. It therefore showed the multiple ways that, instead of the supposed contextualisation of knowledge that Mode 2 purported as the new ways of producing expertise, it was indeed the construction of universal numerical narratives that gave heart and soul to the global education governance game. The following chapter will contest another key principle of the Mode 2 paradigm, namely interdisciplinarity. Through an exploration of the developments that led to the establishment of the SDG4, it will show that rather than interdisciplinary knowledge; it was mono-disciplinarity that led the way and especially the production of a particular kind of knowledge: that is, economic knowledge and its very specific disciplinary and epistemological assumptions.

REFERENCES

- Alonso, W., & Starr, P. (Eds.). (1987). *The politics of numbers*. Russell Sage Foundation and Planning and Duke University Press.
- Barnett, N. M., & Finnemore, M. (1999). The politics, power, and pathologies of international organizations. *International Organization*, 53, 699–732. <https://doi.org/10.1162/002081899551048>

- Barry, A. (2012). Political situations: Knowledge controversies in transnational governance. *Critical Policy Studies*, 6(3), 324–336.
- Baumgartner, F. R., & Jones, B. D. (1993). *Agendas and instability in American politics*. University of Illinois Press.
- Bhuta, N. (2012). Governmentalizing sovereignty: Indexes of state fragility and the calculability of political order. In B. Kingsbury, S. Merry, & K. Davis (Eds.), *Indicators as technologies of global governance*. Oxford University Press.
- Bogdandy, A., Dann, P., & Goldmann, M. (2008). Developing the publicness of public international law: Towards a legal framework for global governance activities. *German Law Journal*, 9, 1375–1400.
- Bourdieu, P. (1993). *The field of cultural production*. Polity Press.
- Bowker, G., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. MIT Press.
- Brosig, M. (2011). Overlap and Interplay between international organisations: Theories and approaches. *South African Journal of International Affairs*, 18(2), 147–167.
- Clarke, J., & Ozga, J. (2011, July 4–6). *Governing by inspection? Comparing school inspection in Scotland and England* (Paper for Social Policy Association conference, University of Lincoln). http://www.social-policy.org.uk/lincoln2011/Clarke_Ozga.pdf. 9 November 2011.
- Cox, R. (1992). Multilateralism and world order. *Review of International Studies*, 18(2), 161–180.
- Cox, R. (1996). The executive head: An essay on leadership in international organization. In R. Cox (Ed.), *Approaches to world order*. Cambridge University Press.
- Cram, L. (2011). The importance of the temporal dimension: New modes of governance as a tool of government. *Journal of European Public Policy*, 18(5), 636–653. <https://doi.org/10.1080/13501763.2011.586793>
- Desrosieres, A. (1998). *The politics of large numbers: A history of statistical reasoning*. Harvard University Press.
- Djelic, M.-L., & Sahlin-Andersson, K. (Eds.). (2006). *Transnational governance: Institutional dynamics of regulation*. Cambridge University Press.
- Espeland, W., & Stevens, M. L. (2008). A sociology of quantification. *Archives Europeennes de Sociologie*, 49(3), 401–436.
- Fischer, F., & Forester, J. (Eds.). (1993). *The argumentative turn in policy analysis and planning*. Duke University Press.
- Fougner, T. (2008). Neoliberal governance of states: The role of competitiveness indexing and benchmarking. *Millennium: Journal of International Studies*, 37(2), 303–326.
- Galaskiewicz, J. (1985). Interorganizational relations. *Annual Review of Sociology*, 11, 281–304.

- Gehring, T., & Oberthür, S. (2009). The causal mechanisms of interaction between international institutions. *European Journal of International Relations*, 15(1), 125–156.
- Global Education for all Meeting/UNESCO. (2014). *2014 GEM final statement: The Muscat agreement*. UNESCO.
- Grek, S. (2009). Governing by numbers: The PISA effect in Europe. *Journal of Education Policy*, 24(1), 23–37.
- Grek, S. (2012). (2012) What PISA knows and can do: Studying the role of national actors in the making of PISA. *European Educational Research Journal*, 11(2), 243–254.
- Grek, S. (2014). OECD as a site of coproduction: European education governance and the new politics of policy mobilization. *Critical Policy Studies*. <https://doi.org/10.1080/19460171.2013.862503>
- Grek, S. (2020). Prophets, saviours and saints: Symbolic governance and the rise of a transnational metrological field. *International Review of Education*, 66, 139–166.
- Grek, S., & Lingard, B. (2023 accepted). Education. In F. de Francesco & C. Radaelli (Eds.), *The Elgar companion to the OECD*. Elgar.
- Haas, P. (Ed.). (1992). Epistemic communities. *International Organization*, 46(1, Special issue).
- Hacking, I. (1990). *The taming of chance*. Cambridge University Press.
- Hacking, I. (2007). Kinds of people: Moving targets. *Proceedings of the British Academy*, 151, 285–318.
- Hall, J. (2009). *The global project measuring the progress of societies: A toolkit for practitioners*. Presentation from the 3rd OECD World Forum, 27–30 October, Busan.
- Hansen, K., & Porter, T. (2012). What do numbers do in transnational governance? *International Political Sociology*, 6, 409–426.
- Jullien, B., & Smith, A. (2010). Conceptualising the role of politics in the economy: Industries and their Institutionalizations. *Review of International Political Economy*, First published on 19 October 2010 (iFirst). <https://doi.org/10.1080/09692291003723615>
- Kaplan, T. J. (1986). The narrative structure of policy analysis. *Journal of Policy Analysis and Management*, 5(4), 761–778.
- Lagroye, J. (1997). *Sociologie politique*. Dalloz-Presses de la FNSP.
- Lampland, M., & Starr, S. L. (Eds.). (2009). *Standards and their stories. How quantifying, classifying and formalizing practices shape everyday life*. Cornell University Press.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.

- Lingard, B., Rawolle, S., & Taylor, S. (2005). Globalising policy sociology in education: Working with Bourdieu. *Journal of Education Policy*, 20(6), 759–777.
- Martens, K. (2007). How to become an influential actor: The comparative turn in European educational policy. In K. Martens, A. Rusconi, & K. Leuze (Eds.), *New arenas of education*. Palgrave Macmillan.
- Merry, S. E. (2011). Measuring the world indicators, human rights, and global governance. *Current Anthropology*, 52, S83–S95.
- Nóvoa, A., & Yariv-Mashal, T. (2003). Comparative research in education: A mode of governance or a historical journey? *Comparative Education*, 39(4), 423–439.
- Nóvoa, A. (2002). Ways of thinking about education in Europe. In A. Nóvoa & M. Lawn (Eds.), *Fabricating Europe—The formation of an education space (António)* (pp. 131–155). Kluwer Academic Publishers.
- OECD. (2003). *What is PISA*. <http://www.pisa.oecd.org/dataoecd/51/27/37474503.pdf>
- Oliver, C. (1990). Determinants of interorganizational relationships: Integration and future directions. *Academy of Management Review*, 15, 2.
- Ozga, J., & Lingard, B. (2007). Globalisation, education policy and politics. In B. Lingard & J. Ozga (Eds.), *The Routledge Falmer reader in education policy and politics* (pp. 65–82). Routledge Falmer.
- Palan, R. (2006). *The offshore world: Sovereign markets, virtual places, and nomad millionaires*. Ithaca.
- Porter, T. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.
- Porter, T., & Webb, M. (2004). *The role of the OECD in the orchestration of global knowledge networks* (Paper prepared for the international studies association annual meeting, March, Montreal).
- Power, M. (1997). *The audit society: Rituals of verification*. Oxford University Press.
- Raustiala, K., & Victor, D. G. (2004). The regime complex for plant genetic resources. *International Organization*, 59(2), 277–298.
- Richardson, J. J. (2001). *European Union: Power and policy making* (2nd ed.). Routledge.
- Ricoeur, P. (1984). *Time and narrative*. University of Chicago Press.
- Ringel-Bickelmeier, C., & Ringel, M. (2010). Knowledge management in international organisations. *Journal of Knowledge Management*, 14(4), 524–539.
- Rittberger, V., & Zangl, B. (2006). *International organization: Polity, politics and policies*. Palgrave Macmillan.
- Rose, N. (1999). *Powers of freedom: Reframing political thought*. Cambridge University Press.

- Saetnan, A. R., Lomell, H. D., & Hammel, S. (Eds.). (2011). *The mutual construction of statistics and society*. London: Routledge.
- Sauder, M., & Espeland, W. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review*, 74, 63–82.
- Shanks, C., Jacobson, H. K., & Kaplan, J. (1996). Inertia and change in the constellation of intergovernmental organizations, 1981–1992. *International Organization*, 50(4), 593–627.
- Shapin, S., & Schaffer, S. (1985). *Leviathan and the air-pump: Hobbes, Boyle and the experimental life*. Princeton University Press.
- Shore, C., & Wright, S. (2015). Audit culture revisited: Ratings, rankings and the reassembling of society. *Current Anthropology*, 56(3), 421–444.
- Smith, A. (2009). *Studying the Government of the EU: The promise of political sociology*. University of Edinburgh.
- Strathern, M. (Ed.). (2000). *Audit cultures: Anthropological studies in accountability, ethics, and the academy*. Routledge.
- UN. (2023). *Sustainable Development Goal 4*. At <https://sdgs.un.org/goals/goal4> [16/11/2023]
- UNESCO. (2016). *Incheon declaration and framework for action for the implementation of sustainable development goal 4*. UNESCO.
- Van de Ven, A. (1976). On the nature, formation, and maintenance of relations among organizations. *Academy of Management Review*, 1(4), 24–36.
- Weick, K. E. (1995). *Sensemaking in organisations*. Sage.
- Wynne, B. (2003). Seasick on the third wave: Subverting the hegemony of propositionalism. *Social Studies of Science*, 33(3), 401–417.
- Young, O. (1996). Institutional linkages in international society: Polar perspectives. *Global Governance*, 2, 1–24.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





The Rise of Mono-disciplinarity: Learning, Economics and the Production of Non-knowledge

1 INTRODUCTION

Although the move from Mode 1 to Mode 2 knowledge production was meant to herald a new era of multidisciplinary in understanding, studying and dealing with social problems, the experience of the ways expert knowledge is produced and governed is far from the ideal evolutions that such schematic representations tend to offer. Indeed, despite the intractability of the problems at hand, global education governance has not been informed by the vision and multiplicity of perspectives that arise when different forms of knowledge come together. Instead, the datafication of education problems severely restricted the scope of knowledge production. Even worse, the quantification of education problems quickly led to their economisation, as economists of education rose to positions of authority in major IOs and hence their concerns around issues of input/output mechanisms, efficiency of education systems and cost/benefit analyses began to dominate the debate of how to bring improvement and reduce inequities. Thus, instead of Mode 2's proclaimed multidisciplinary epistemological perspective, what is to be observed in the field of education—and arguably more broadly—is that the discipline of economics has emerged as the great unifier that brings together actors, narratives and policy solutions.

Therefore, this chapter will argue that the complexity of policy problems led not necessarily to the rise of interdisciplinarity, but rather in

mono-disciplinarity; as a result, certain kinds of knowledge were privileged, while others were silenced in the process. Since the 1990s, there has been a slow but concerted effort to use quantification as an instrument of the *economisation* of education discourse and practice as the single, universal language of global education policy: as we will see, the economisation of global education policy, with its emphasis on comparability, efficiency and cost-effectiveness, enabled the communication between different disparate groups and became the language and policy of choice for education policymakers around the world.

The chapter uses the notion of epistemic infrastructures (Tichenor et al., 2022) in order to empirically chart this process as having happened at three levels: (1) economisation that occurred through the central positioning of economists within the education inequity debate; (2) through the expansion of an economic/instrumental way of thinking about education; (3) finally, economisation occurred through the increasingly central role of international organisations whose primary remit centres around economic growth concerns (such as the OECD and the World Bank). What we observe, as a result, is the construction of economic knowledge in education, at the expense of other perspectives, and thus, as we will see, the simultaneous production of knowledge and non-knowledge as part of the process. This dichotomy, namely between the production of certain types of data at the expense of others will be discussed at the final sections of the chapter, in an effort to understand the effects of mono-disciplinarity in global education governance.

2 THE CONTOURS OF SDG4: THEORISING WITH EPISTEMIC INFRASTRUCTURES

As already discussed in the previous chapter, the World Education Forum (WEF) was celebrated in Incheon, the Republic of Korea, in May 2015, with the participation of over 1500 people, including 120 Ministers of Education and representatives from a wide range of international governmental and non-governmental organisations. The event at Incheon represented a milestone in the history of UNESCO summitry, a long trajectory of large education conferences that demanded fair, free and quality education for all. Similar to others prior to it, the main product of WEF 2015 was the so-called Incheon Declaration, along with the Framework for Action adopted by UNESCO Member States a few months later,

in November 2015. In conjunction, both documents established an ambitious and highly aspirational education agenda for the period 2015–2030 and condensed in the overarching goal to ‘ensure inclusive and equitable quality education and promote lifelong learning opportunities for all’ and a number of associated targets; this is the Sustainable Development Goal 4 (SDG4) (UNESCO, 2016).

Indeed, the SDG4 is one of the 17 Sustainable Development Goals (SDGs) that are ‘integrated and indivisible and balance the three dimensions of sustainable development: the economic, social and environmental’ (UN, 2015, p. 5). According to the UN, ‘they result from what is arguably the most inclusive process of consultation in the history of the United Nations, reflective of substantive input from all sectors of society, all actors of the international community and all parts of the world’ (UNESCO, 2017, p. 4). Indeed, as will be shown here, it is precisely this inclusive and participatory governance model that became key in the formation of many aspects of the SDG4 agenda and its implementation. As a programmatic document oriented at nurturing and securing a form of collective commitment towards a shared set of aspirations, the new agenda builds on a well-established tradition of consultation and collaboration that has come to be recognised as a characteristic of the UN system. What is interesting—and will be discussed later in this chapter—is that, despite the proclaimed collective and broad set of aspirations, certain kinds of data production for specific indicators remain dominant, at the expense of a focus on others.

Thus, this chapter traces the development of the epistemic infrastructure of the SDG4 in order to show the ways that the incremental build-up of the discourse, technical expertise and, given this apparent universality of the SDG agenda, the fragile but necessary actor 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. The shift entailed the prioritisation of an emphasis 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 or length of the school year, the pendulum shifted to a greater interest in decontextualised, applied knowledge, measured in real-life contexts. 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; 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, alongside other key venues (one of them being OECD’s PISA, as will be discussed further on) the SDG4 became a prime site of the production of this radical reconceptualization of education measurement and policy with implications across the world.

Indeed, the complexity and length of the SDG4 process render the painting of a comprehensive picture of all related events and actors as a futile endeavour. A focused analysis of the production of the SDG4, viewed through the lens of the notion of ‘epistemic infrastructures’, allows for a close-up on the interdependency of materialities, technologies, individual actors and organisations that participated in its making. Indeed, the paper adopts the definition of an epistemic infrastructure as the ‘complex interplay of material, techno-political and organisational structures within which (statistical) knowledge is produced, disseminated and translated into global public policy’ (Tichenor et al., 2022).

Earlier literature on infrastructure studies (Star, 1999; Winner, 1986) highlighted their invisibility; infrastructures were seen as comprised by social, material and technological elements that are interdependent and flow seamlessly into one another, facilitating the unobstructed move of numbers, people, goods and ideas in the production of new ways of measuring, viewing and living in this world. However, in contrast to the neat accounts of global education reforms flowing top-down, the SDG4 has never been the perfect invisible infrastructure, moving ideas and practices from some imaginary ‘centre of calculation’ (Latour, 1987) to the periphery. Instead, long before its inception, it has been a site of conflict and contestation, a space where relationships break-down and—more often than not—metrics fail. Since the idea of metric ‘failure’ might have normative connotations, it needs to be clarified that I examine ‘failing metrics’ as those that lose their policy momentum, by increasingly being perceived by the policy, expert and professional communities as irrelevant or even misleading; ultimately, their continued measurement is seen as having detrimental, rather than positive effects on the policy arenas they are meant to contribute. Such failings can be either real or manufactured, yet the outcome is the same: the failure of achieving global goals (irrespective of whether they are misplaced or, in fact, unattainable in the first

place) sparks quests for improved metrics that will excite, persuade and ‘stick’ anew (Bandola-Gill, 2020). Yet, despite such perceived failures, it is the infrastructure’s break-down that fuels its growth and expansion. As this chapter will show, the paradigmatic shift from the policy focus on schooling to learning happened through the concerted efforts to discredit certain kinds of knowledge production, in favour of others that were seen as linking education a lot closer to economic prosperity; that is, economic knowledge.

The policy prioritisation of learning and its associated outcomes is not a novel topic in education research. Although there has been scholarship on the discursive expansion of the language of learning outcomes and skills (Klees et al., 2019), as well as some critical literature on the validity and robustness of the new learning metrics (Benavot & Smith, 2020), and on their effects on global education policy reforms (Mundy et al., 2016), the chapter discusses the *entanglement* of materialities, discourses, ideas and practices into the building of a new epistemic infrastructure that has prioritised the dominance of the discipline of economics in education governance globally.

Indeed, these entanglements have allowed a plethora of contestations to unfold: one of the most prominent ones is the emphasis on some indicators versus others, as well as the issue of the democratic decision-making process. After a brief overview of the intellectual terrain on infrastructures and some methodological considerations (Sect. 2), the following section (Sect. 3) will discuss the history of the shift of education discourse from the measurement of inputs to the dominance of measuring skills and outcomes. In particular, I will discuss the ways in which some powerful actors prepared the ground for a move away from the measurement of schooling (through measuring access and completion) to learning (through the measurement of literacy skills). The primary means of facilitating this change was through presenting the MDG education targets as misleading and thus as ‘failing’ metrics; the mobilisation of new evidence and a ‘killer’ number (Stevens, 2011) was used in order to create the space for contestation and change. The building of a discourse of the economic versus the wider social benefits of education, alongside the production of new metrics to replace the old ones, became a vital mix and thus the building block in the construction of the infrastructure of the SDG4. In addition, I will highlight the importance of the *temporal* dimension in the building of epistemic infrastructures, in terms of first, their temporal discursive framing of ‘past failures/ current crisis/ future projections’,

as well as in relation to the slow, step-by-step build-up of the measurement infrastructure in order to gather steam, create the evidence, build a support base and thus have greater policy influence.

Section 4 will then move on to the analysis of the workings of the Technical Advisory (later Cooperation) Group, in charge of the development of some of the SDG4 indicators. The section will show how the TAG/TCG began its work in 2014 primarily as a group of expert IO statisticians and later expanded into a much larger—and with a different function—grouping that included country and civil society representatives, all in the name of democratising the measurement agenda and process. Thus, beginning with the small, highly technical and elitist group in 2014, we observe how the slow building of a much larger infrastructure of actors and materialities came together to support, prop up and legitimise the work of the production of numbers. Thus, this section will focus more on the *spatial* features of the infrastructure, as it expanded across contexts and fields of practice, to include a much wider actor membership and achieve greater coordination across the local, national and global levels.

Section 5 will discuss the infrastructural qualities of meetings of the SDG4, by showing how, instead of a seamless flow of coordination and cooperation, it was failing metrics and the continued break-down of the proceedings that both acted as generative forces that ensured its continuity and growth. These meetings that bring together a range of actors, from the local to the international levels are, as I will show below, those slow and convoluted processes that ‘wicked’ problems (Guy Peters, 2017) are discussed and a range of possible monitoring solutions agreed upon. As this chapter shows, the process of the rise of mono-disciplinarity in education requires not only the co-construction of specific kinds of knowledge by the relevant IOs, but also another significant function of theirs: that of the making of ignorance, or as we prefer to call it, non-knowledge. The social production of non-knowledge is a necessary pre-condition for reaching agreement about what kind of knowledge will be pursued in order to achieve a minimum consensus, so as to ensure ‘buy-in’ but also maintain actors’ own interests, values and positions intact (Grek, 2020). Thus, the construction of non-knowledge is an essential part of the measurement process: rather than the opposite of knowledge, however, or its reading as a binary, here it is viewed as a symbiotic relationship, necessary for balancing out and achieving some kind of constant equilibrium—and hence movement—of the metrological field.

Finally, in terms of methods, the chapter is built on three main sources of data: first, the discourse analysis of documents relating to the production of the SDG4, as well as materials that predated it. CDA is a particularly apt method for the analysis of the making of infrastructures because it sees text as a key aspect of how certain understandings of the world are shaped and perpetuated by practices of infrastructuring (Meyer, 2001; Wodak & Fairclough, 1995). Hence, the analysis of these documents is useful for, on the one hand, showing what is technically possible, while on the other, explaining what the principles and perspectives of those participating in the production of the infrastructure are.

Second, the chapter's empirical analysis is based on twenty in-depth interviews with key actors of international organisations and the civil society. Finally, the social network analysis component focused on an exploration of the role of the SDG4 meetings and the alliances and connections they generated. The combination of these methods allowed for a study of the discursive meaning produced by relevant IO and research reports. Interviews gave me an insight into the experience, views, positionings and choices of the key actors that participated in the infrastructure. Lastly, social network analysis, focusing on the two main indicator technical groups, explored their meetings as the key stabilising moment when negotiations achieved the desired *pax romana* before disagreement and conflict unravelled again. Thus, the research design offered the capacity to study different elements of the infrastructure, their entanglements, effects and the ways certain kinds of knowledge production dominated over others.

3 THE RISE OF INFRASTRUCTURES: VOGUE, VAGUE OR 'REALLY USEFUL KNOWLEDGE'¹?

'Infrastructures are conceptually unruly' (2013, p. 329), Brian Larkin wrote, and there could not have been a more accurate description for the varied application of the term. In fact, it is precisely the conceptual plasticity *and* the focus on materiality that has made infrastructures such a popular concept in social theory. Nonetheless, they have not always been as vogue as they are today: in fact, it was only in the mid-1990s when

¹ The term 'really useful knowledge' is derived from radical education thought of the nineteenth century; it was supporting a critical understanding of self and society; it was knowledge meant 'to set you free'.

Geoffrey Bowker (1995) first pointed towards the materiality of infrastructures as a way of understanding their function and effects. Bowker saw infrastructures as largely invisible backdrops to social action and thus analytically not penetrable; he therefore proposed the notion of ‘infrastructural inversion’, as a way of breaking the invisibility and flow of the infrastructure. Infrastructural inversion (Bowker, 1995) was about making the invisible visible, through a focus on material relations and the ways they reconfigure how we know and live in the world.

Similarly, in 1996, Susan Leigh Star and Karen Ruhleder saw invisibility as a key quality of infrastructural systems. Nonetheless, they also identified the seamless flow of the infrastructure as a fragile achievement that was prone to break-down and failure (Star & Ruhleder, 1996). The invisibility/visibility conundrum was further discussed by Larkin (2013), who suggested that infrastructures can be invisible but can also become a spectacle, and thus depend on their visibility for their success. However, following Larkin, even when an infrastructure is open, visible and ready to be experienced, what is there to see? According to Harvey et al.,

Provisionally, and minimally, we might say that we are dealing with technologically mediated, dynamic forms that continuously produce and transform sociotechnical relations. That is, infrastructures are extended material assemblages that generate effects and structure social relations, either through engineered (i.e. planned and purposefully crafted) or non-engineered (i.e. unplanned and emergent) activities. (2017, p. 5)

This analysis contributes to the literature on infrastructures, by showing the particularities of the mix of materials, practices and meanings in the making of measurement agendas, such as the SDGs. Given the centrality of knowledge and data production in global governance, the concept of ‘epistemic infrastructures’ (Tichenor et al., 2022) is particularly apt, since it advances the analytical purchase of the—STS-primarily informed—concept to bring it much closer to policy theory and practice. In particular, as the chapter shows, both the flow and the failures, the unlikely alliances and the clashes, did not only facilitate the production of a system of measurement and a particular way of naming and understanding educational realities in the twenty-first century. They also brought about a much more fundamental and—as it appears—permanent policy shift: this was the dominance of the economic paradigm in education measurement, practice and values. The move away from the measurement and thus

prioritisation of educational inputs (numbers of teachers, school facilities, financial support and others) towards the measurement of outputs (learning outcomes, test results, skills and competencies) did not merely take place at the discursive level, or the measurement one. Neither has it only been circulated and promoted among organisations and actors, experts and professionals, that work in the field of education. Rather, it produced a monodisciplinary dominance in the education policy field that has had dramatic consequences in the way education policies at the country level are made (Verger et al., 2019). The intention of the chapter is to utilise the three orders of the epistemic infrastructure (the materialities, the interdependencies and the paradigmatic shifts) in order to place emphasis on the role of the discipline of economics for producing knowledge for policy.

4 FROM SCHOOLING TO LEARNING: THE INCREMENTAL BUILDING OF AN INFRASTRUCTURAL BASE, 2006–2013

The discursive and logical shift 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 decontextualizing 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 3 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). In many senses, OECD PISA became the flagship international comparative test that shifted the focus of education policymakers to outputs, rather than inputs, and to learning rather than schooling. The significance of PISA data is undisputable, given

that European education governance became dependent on it, in order to—for the first time ever—create indicators and benchmarks to measure education performance in EU member states—what was called the Lisbon agenda (Lawn & Grek, 2012).

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 school 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 seminal 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 lead 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 example of the failed MDGs is an excellent illustration of the core argument of this paper in regard to the power of metrics not only to influence policy direction, but also in fact to be the space where policy work is done: it was the production of new metrics by education economists that pushed for the idea that previous metrics had failed. And it is precisely the perceived failure of the MDGs that created the new space for contestation around which new metrics (and thus policy priorities) should replace them. The

materiality of data, reports and meetings intersected with the work of specific expert organisations and actors and led to a substantial policy shift. As I discussed, these expert actors were international organisations with a very explicit mission and objective: that is, to increase economic growth and development.

Indeed, the arguments developed by education economists at 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 of 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).

Indeed, as the decade progressed and the end of the MDG timeframe was drawing to a close, we can observe a much more concerted effort to change not only 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 targets. The key protagonist in this new era was not the World Bank (though it was always supporting at 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 economists 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 economies and 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 organization that included a wide range of actors not only from the international organisations’ expert world, but also 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; it also offered a way to break away from the dominance of economic thinking and allow a broader conversation. 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.

5 FROM IOS' ADVISORY TO COOPERATIVE ROLE: BROKERAGE AND COLLABORATION

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, economists extended an olive branch to academics, the civil society and professional organisations that perceived the learning focus as reductionist and as reflecting merely the economic lens of the Bank's ideological positioning. Additionally, UIS' leadership (and not the World Bank's, for example) 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 on-going open consultations led by regional leaders. Very quickly, what was a small, rather swift and efficient technical team of IO education economists (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 on the footsteps from 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. Wider legitimacy was sought and gained through the expansion of the measurement infrastructure into an epistemic one: one that became legitimate and dominant through its active involvement of actors from across sectors and countries. Even though the involvement of the majority of these actors was generally passive, the language of the new indicators became the new *episteme*: that is, a way of knowing, describing and communicating about the world that was not merely about the craft of numbers but involved the production of a new governing paradigm: that of the dominance of the mono-disciplinarity of economics in global education governance.

6 CONSTRUCTING NON-KNOWLEDGE: MONO-DISCIPLINARITY AND THE SILENCING OF ALTERNATIVE PERSPECTIVES

Indeed, as we saw in the previous section, the open, inclusive and participatory nature of the consultative process facilitated by UNESCO and the EFA architecture was in many ways unprecedented, and the openly negotiated and improvisatory character of the SDG debate contrasted with the technocratic origins of the MDGs (cf. Fukuda-Parr & McNeill, 2019).

In many ways, it is precisely this open debate and the participatory nature of the SDG governing architecture that has allowed a plethora of contestations to unfold: one of the most prominent ones is the large emphasis on some indicators (especially those that measure performance in literacy and mathematics) that comprise goal 4 versus others. Table 1 offers a useful overview of the different indicators in goal 4.

Although the development of SDG4 has been described as ‘arguably the most inclusive process of consultation in the history of the United Nations’ (Naidoo, 2016), this was not matched by the making of the relevant indicators to measure the ambitions (McGrath & Nolan, 2016; Smith, 2019). As discussed in previous sections, the process became quite technical from the start. Statisticians and their considerations for valid and robust data took hold of the process and most non-statistical knowledge

Table 1 The SDG4 indicators are as follows: 4.1.1 on reading and maths proficiency; 4.2.1/2 on early childhood; 4.3.1 on VET; 4.4.1 on ICT skills; 4.5.1 on gender equality; 4.6.1 on adult literacy and numeracy; 4.7.1 on global citizenship and sustainable development. Available at <http://uis.unesco.org/sites/default/files/documents/11-global-indicators-sdg4-cheat-sheet-2018-en.pdf>

11 Global Indicators for SDG 4		SUSTAINABLE DEVELOPMENT GOALS		
TARGETS	<p>4.1.1 #learning #math #proficiency</p> <p>Proportion of children and young people (a) in Grade 2 or 3, (b) at the end of primary education, and (c) at the end of lower secondary education achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex.</p> <p>CONCEPT: Minimum proficiency level (MPL) is the benchmark for the minimum level of skills that children and young people etc. measured through learning assessments. Currently, there is limited comparability of data from different learning assessments.</p> <p>DATA SOURCES: Cross-national learning assessments (PASEC, PIRLS, PISA, SACMEQ, TERCE, TIMSS) and national assessments.</p>	<p>4.2.1 #EarlyChildhood #ChildDevelopment</p> <p>Proportion of children under 5 years of age who are developmentally on track in health, learning and psychosocial well-being, by sex.</p> <p>CONCEPT: There is not yet a globally-accepted definition of developmental milestones for children under 5 years of age. ECDF defines 'on track' as children aged 3 to 4 years who are developmentally on track in at least three of these four domains: literacy-numeracy, physical, socio-emotional and learning.</p> <p>DATA SOURCES: Measures to capture children's early childhood experiences have been used in multiple countries in projects such as the UNICEF Multiple Indicator Cluster Surveys (MICS).</p>	<p>4.2.2 #EarlyChildhood #ECE</p> <p>Participation rate in organized learning (one year before the official primary entry age) is the age at which children are obliged to start primary education according to national legislation or policies.</p> <p><i>Part rate = $\frac{\text{Enrollment (ICD2) or 1} + \dots}{\text{Population (one year age)}}$</i></p> <p>DATA SOURCES: Enrollment data reported by Ministries of Education or national statistical offices and population attendance data from household surveys and censuses.</p>	<p>4.3.1 #formaleducation #training</p> <p>Participation rate of youth and adults in formal, non-formal education and training in the previous 12 months, by sex.</p> <p>CONCEPT: Methodology related to participation in formal education and training in the previous 12 months, most courses (ISCED 2011). However, methodology for measuring participation in non-formal education and training varies substantially across the globe.</p> <p>DATA SOURCES: At the international level, surveys such as the European AEL, PIACC from OECD and the LO SMITS.</p>
	<p>4.4.1 #skills #ICT</p> <p>Proportion of youth and adults with information and communications technology (ICT) skills, by type of skill.</p> <p>CONCEPT: The indicator measures ICT skills based on the number of people who report having undertaken certain computer-related activities in a given time period.</p> <p>DATA SOURCES: Activities are categorized by Eurostat and adopted by the International Telecommunications Union (ITU). Both organizations coordinate national surveys to collect this information.</p>	<p>4.5.1 #equity #parity</p> <p>Parity indices (female:male, rural:urban, bottom/top wealth quintiles and others such as disability status, indigenous peoples and vulnerable groups) of indicators of educational attainment in all education indicators on this list that can be disaggregated.</p> <p>CONCEPT: The indices represent the ratio of the indicator value for one group to that of the other. Typically, the likely more disadvantaged group is in the numerator. A value of exactly 1 indicates parity between the two groups.</p> <p>DATA SOURCES: Same as for underlying indicators.</p>	<p>4.6.1 #proficiency #literacy</p> <p>Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex.</p> <p>CONCEPT: Level of proficiency in the benchmark of basic knowledge in a domain measured through learning assessments. So far, there are no common standards validated by the international community or countries.</p> <p>DATA SOURCES: Skills assessment surveys of the adult population (e.g. PIACC, STEP, LAMP and national literacy and numeracy surveys).</p>	<p>4.7.1 #globalcitizenship #sustainabledevelopment</p> <p>Extent to which (i) global citizenship education and (ii) education for sustainable development, including gender equality and human rights, is integrated into national education, and (iii) student assessments.</p> <p>CONCEPT: It seeks to measure the quantity and quality of country inputs, as well as whether the quality of OECD and ESD provision is adequate to fulfil their transformational potential.</p> <p>DATA SOURCE: UNESCO consultation on the implementation of the Sustainable Development Goals, Cooperation and Peace and Education relating to Human Rights and Fundamental Freedoms.</p>
	<p>4.a.1 #LearningEnvironment #facilities</p> <p>Proportion of schools with access to: (a) electricity, (b) Internet for pedagogical purposes, (c) computers for pedagogical purposes, (d) adapted infrastructure and materials for students with disabilities, (e) basic drinking water, (f) single-sex basic sanitation facilities, and (g) basic handwashing facilities (a per WASH indicator definition).</p> <p>DATA SOURCES: Administrative data from schools and other providers of education or training.</p>	<p>4.b.1 #scholarships #developmentaid</p> <p>Volume of official development assistance flows for scholarships by sector and type of study.</p> <p>DATA SOURCES: Administrative data on disbursement of official development assistance from OECD's Development Assistance Committee.</p>	<p>4.6.1 #proficiency #literacy</p> <p>Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex.</p> <p>CONCEPT: Level of proficiency in the benchmark of basic knowledge in a domain measured through learning assessments. So far, there are no common standards validated by the international community or countries.</p> <p>DATA SOURCES: Skills assessment surveys of the adult population (e.g. PIACC, STEP, LAMP and national literacy and numeracy surveys).</p>	<p>4.7.1 #globalcitizenship #sustainabledevelopment</p> <p>Extent to which (i) global citizenship education and (ii) education for sustainable development, including gender equality and human rights, is integrated into national education, and (iii) student assessments.</p> <p>CONCEPT: It seeks to measure the quantity and quality of country inputs, as well as whether the quality of OECD and ESD provision is adequate to fulfil their transformational potential.</p> <p>DATA SOURCE: UNESCO consultation on the implementation of the Sustainable Development Goals, Cooperation and Peace and Education relating to Human Rights and Fundamental Freedoms.</p>
	<p>4.a.1 #LearningEnvironment #facilities</p> <p>Proportion of schools with access to: (a) electricity, (b) Internet for pedagogical purposes, (c) computers for pedagogical purposes, (d) adapted infrastructure and materials for students with disabilities, (e) basic drinking water, (f) single-sex basic sanitation facilities, and (g) basic handwashing facilities (a per WASH indicator definition).</p> <p>DATA SOURCES: Administrative data from schools and other providers of education or training.</p>	<p>4.c.1 #teachers</p> <p>Proportion of teachers in (a) pre-primary education, (b) primary education, (c) lower secondary education, and (d) upper secondary education who have received at least the minimum organized teacher training (e.g. pedagogical training) pre-service or in-service required for teaching at the relevant level in a given country, by sex.</p> <p>DATA SOURCES: Administrtive data from schools and other organized learning centres.</p>	<p>4.6.1 #proficiency #literacy</p> <p>Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex.</p> <p>CONCEPT: Level of proficiency in the benchmark of basic knowledge in a domain measured through learning assessments. So far, there are no common standards validated by the international community or countries.</p> <p>DATA SOURCES: Skills assessment surveys of the adult population (e.g. PIACC, STEP, LAMP and national literacy and numeracy surveys).</p>	<p>4.7.1 #globalcitizenship #sustainabledevelopment</p> <p>Extent to which (i) global citizenship education and (ii) education for sustainable development, including gender equality and human rights, is integrated into national education, and (iii) student assessments.</p> <p>CONCEPT: It seeks to measure the quantity and quality of country inputs, as well as whether the quality of OECD and ESD provision is adequate to fulfil their transformational potential.</p> <p>DATA SOURCE: UNESCO consultation on the implementation of the Sustainable Development Goals, Cooperation and Peace and Education relating to Human Rights and Fundamental Freedoms.</p>
	MEANS OF IMPLEMENTATION			

was excluded. This was not however the only omission; perhaps the more significant one took place when there was an early decision upon some indicators which would be considered ‘global’ versus those that were relegated to the description of ‘thematic’ (Smith, 2019). This was a key moment, since,

While global indicators are universally applied and expected to be reported by all countries, thematic indicators are considered voluntary. Therefore, the majority of resources in indicator creation, monitoring, reporting and state action will focus on the global indicators while thematic indicators are not taken into account in the UN’s annual SDG report. (Smith, 2019, p. 3)

The pendulum had already swung. Although target 4.1 was promising that ‘by 2030, ensure that all girls and boys complete free, equitable and quality primary education leading to relevant and effective learning outcomes’, the 4.1.1 global indicator that came to be associated with it was much more limited in reporting on ‘quality’ only, whereas free and equitable education were downgraded to ‘thematic’, if they were even considered at all (King, 2017). In other words, the production of certain knowledge was privileged over others; this of course was done (and is always done) on the basis of the methodological robustness and validity of the exercise.

Indeed, many of the interviewees that METRO examined, suggested that the fundamental problem of the SDGs lies in the fact that it began the process by setting the ambitions and establishing the goals, rather than checking whether there was enough data or the right methodologies to monitor them. Nevertheless, the (limited arguably) resources that were put in the process were invested in indicators that were already backed up with significant statistical evidence. The strategic choice to construct non-knowledge by emphasising some indicators versus others becomes even more evident in the tensions that the negotiations around indicator 4.7 as created. As Antonia Wulff, from Education International, contends,

The expert group in charge of the SDG indicators rejected the proposed measurement strategy for target 4.7 on education for sustainable development, human rights and global citizenship... Education International is generally concerned about the slow progress made on key indicators and, importantly, the large disparity in the time, effort and resources put into

developing 4.7 indicators as opposed to the learning outcomes under target 4.1. We are impatient to move forward. (Wulff, 2018)

Although limitations of space in the present chapter do not allow for a more extensive empirical analysis of the privileging of certain kinds of education data production over others, the above example serves as a useful illustration of the making of ‘non-knowledge’; rather than simply an ‘inability-to-know’, strategic decisions were made in relation to which disciplinary perspective was prioritised and took hold. The sociology of quantification has already persuasively discussed how quantification creates visibility, in antithesis to aspects of social life less easy to count. Although a collectively agreed ambition, indicator 4.7 on global citizenship, unless prioritised, measured and backed up with data, will remain a tokenistic representation of those ambitions that turned into ‘goals’ but were then strategically silenced in the process. Thus, at least in the field of transnational performance measurement agendas, the making of any knowledge implies simultaneously the omission of other routes to knowledge, or, in other words, the active production of non-knowledge.

7 DISCUSSION: MONO-DISCIPLINARITY AND QUANTIFICATION

This chapter focused on an analysis of the ways that developments in global education governance since at least the start of the century propped up and legitimised the rise of economics as the dominant epistemology and method for perceiving and solving educational issues of inequity and performance management. Conflicting ideas and interests in this field reveal how epistemic infrastructures, rather than being monolithic blocks, remain fragile and, despite their claims to data and objectivity, are still open to plenty of epistemological and methodological contestation. In the case of the SDG4, it is evident that such contestation and the perceived failure of previous metrics emboldened education economists of major IOs to shift the agenda and move it along. The chapter showed how the perceived failing of the education MDGs (with the use of flagship numbers of emergency, such as the 250 million children not having basic literacy skills) was used as a vehicle to slowly build the dominance of education economics that, although having plenty of ‘concessions’, is now perceived as the dominant disciplinary regime in global education reforms. There were plentiful of circumstances that the disagreement was

such, that a possible break-down seemed almost unavoidable: for example, the reason of the compromise in the drawing of the main parameters of the SDG4 was the real possibility of the exclusion of an education-focused goal, due to the polarisation of the two ‘camps’. Yet, it is precisely the diversity and entanglement of the infrastructure’s social, technical and political elements that sustained and even strengthened the dependence on a single disciplinary field, that of economics, as the only robust and efficient way to measure and evaluate education performance.

Thus, the chapter focuses on the incompleteness and fragility of the infrastructure, alongside the generative power of failing metrics to provide fertile ground for more—and allegedly more precise and truthful—production of data, informed as we have seen both methodologically and epistemologically from a very specific disciplinary perspective. Here the chapter’s focus aligns with Calkins and Rottenburg (2017) in their engagement with ‘infrastructuring as a material-semiotic practice’: although the stable materiality and the techno-scientific dimensions of infrastructural work remain in place, the term is meant to denote the ongoing, continuous nature of infrastructuring as *practice* rather than as a solid, stable space of production. Quantification in epistemic infrastructures becomes the fuel and language of practice, as it brings together ideas and objectivity in one entangled mix. In addition, as we have seen, quantification also lends to the dominance of specific methodological considerations, or at least path dependency with previous data collections, so that there are either crisis calls for needing to change the measurement agenda (as the calls for moving away from education inputs showed) or the dominance of the status quo and the strategic and systematic ignorance of certain types of knowledge. This level of confidence and self-belief, as Fourcade and her colleagues observed, has been a strong characteristic of economists, and perhaps the key quality that distinguishes them from other social scientists:

The fact is that -in some ways true to its philosophical origins- economics is a very moral science after all. Unlike atoms and molecules, the ‘objects’ upon which economists seek to act have a perspective on the world, too. Human life is messy, never to be grasped in its full complexity or shaped according to plan: people act in unanticipated ways; politics makes its own demands; cultures (which economists do not understand well) resist. Thus, the very real success of economists in establishing their professional

dominion also inevitably throws the into the rough and tumble of democratic politics and into a hazardous intimacy with economic, political and administrative power. It takes a lot of self-confidence to put forward decisive expert claims in this context. That confidence is perhaps the greatest achievement of the economics profession -but it is also its most vulnerable trait, its Achilles' heel. (Fourcade et al., 2015, p. 111)

Indeed, one of the main findings of the METRO project (within which the case of the SDG4 was studied) is the changing role of international organisations and their increased confidence to produce not only data but also expert advice on policy directions. Specifically, rather than assuming the expert role of the data producers (therefore asserting their credibility through the production of scientific truth), they have taken a new, brokerage role (Bandola-Gill, 2020; Bandola-Gill et al., 2021; Grek, 2020), working across different institutions and actors and pushing for narratives that link education (alongside other areas of sustainable development) with economic growth and prosperity. Such links are of course not new; since the 1960s and the establishment of many IOs, education was seen as the policy field where interventions and reforms would bring it closer to labour market needs and the production of 'manpower'. Nonetheless, through the increase of the number of actors involved in global governance arrangements, what some have called the 'stakeholderization' of governance, IOs (and especially those who remit involves economic development, such as the World Bank and the OECD) have acquired new prominence and power through monitoring exercises, like the SDGs: as the example of the SDG4 showed the latter create zones of visibility and intervention, while simultaneously produce areas of opaqueness and invisibility.

The chapter has also pointed to two further aspects of the work of infrastructuring that we need to take into account: that is, their temporal and spatial elements. First, starting with the concept of time, any infrastructural investment has a temporal element that is not only evident in the passage of chronological time, but is also palpable in the transformational intent and the promise of a utopian perfectibility; the latter is a key epistemological premise of the discipline of economics, with growth being seen as the goal of achieving everlasting improvement. This promise of an anticipatory better future is central in the work of the SDGs: when it comes to SDG4, it has almost taken a moral dimension and sense of urgency (Grek, 2020), capitalised to either speed up

or slow down the process depending on context. The SDG4 discursive analysis of reports and declarations (see also Chapter 1) shows infrastructural meaning to be produced through gathering past failures and future ideals into an unfolding anticipation in the present, or in other words economics' pursuit of unhindered future growth. The case showed that apart from the anticipatory talk, a certain slowness of time was important in laying down the foundations of the new agenda, avoiding shocks and too sudden changes. Once the groundwork was done, after 2015, we see the process speeding up, coupled with an emphasis on expanding the infrastructure spatially and including a great variety of actors, both geographically, in terms of sectors as well as the ideas and interests that contributed to its production. Again, as already discussed, such expansion of the SDGs in regard to the inclusion of diverse policy actors, including state and non-state ones, was another reason for the prevalence of economics as the quick and accepted disciplinary perspective that would be able to put in line such a wide range of interests and ideas.

Finally, to return to the chapter's earlier discussion, recent years have seen the rise of the sociology of ignorance, a new field of studies that examines the other, less visible side of the politics of constructing knowledge: that is, the politics of ignorance, or as this chapter prefers to call, the politics of 'non-knowledge'. Linsey McGoey has been one of the key advocates of the need for social science to examine 'the mobilisation of ambiguity, the denial of unsettling facts, the realisation that knowing the least amount possible is often the most indispensable tool for managing risks' (McGoey, 2012a, p. 3).

The consideration of the symmetry of knowledge/non-knowledge is of course not new. Socrates insisted that his 'wisdom' was derived by his knowledge of what he didn't know. Philosophically and historically the realisation of the limits of the human knowledge has always been present; nevertheless, our over-emphasis on examining the political uses of knowledge in governing societies has resulted in not engaging nearly enough with non-knowledge. Non-knowledge (or, for others, ignorance) here is not seen as an impediment and obstacle to knowing, but as a *productive force*, that strengthens the role of knowledge and of the knowing subject. For scholars in the field of ignorance studies, we need to investigate non-knowledge as 'regular' rather than 'deviant' (Gross & McGoey, 2015, p. 4). Yet, to date these discussions lack a coherent, agreed-upon nomenclature (Smithson, 2008). Although some scholars use ignorance and non-knowledge interchangeably (e.g. Kleinman &

Suryanarayanan, 2013, p. 495), others distinguish between the two (e.g. Gross, 2012), emphasising the need to avoid the negative connotations that the word ‘ignorance’ implies. Further, there are also scholars who develop taxonomies of different types of ignorance and non-knowledge (e.g. Aradau, 2017; Beck & Wehling, 2012; Gross, 2016).

A review of the literature in the growing field of ignorance studies would be beyond the scope of this chapter. However, the key message that most of this literature appears to agree upon, despite the differences in terminology, is that non-knowledge is productive and not just the negative side of knowledge. Actors may actively try to nurture and preserve ignorance to use it as a resource to advance their interests be it in claiming more funding, denial of responsibility, or assertion of expertise (McGoey, 2012b, p. 555). Importantly, McGoey emphasises that such production and use of non-knowledge may be strategic and deliberate, but not necessarily conscious. Mallard and McGoey go further to propose an epistemological position ‘which asserts as a general maxim that ignorance can be an equally powerful political resource as knowledge’ (2018, p. 3). They suggest that

A second exploration by social scientists of how policymakers, experts and bureaucrats contribute to the production of *soft forms of ignorance* in international affairs... is the literature on the production of indicators, ratings, benchmarks which now circulate everywhere in the world of IOs and global media (Davis, Fisher, Kingsbury, and Merry 2012; Espeland and Sauder 2007; Espeland and Vannebo 2007). As scholars of transparency and auditing practices have long pointed out (cf. Strathern 2000; Power 1997), such indicators help to make policy decisions appear as if they belong to the realm of the certain and unquestionable even when policy options are based on the flimsiest set of factual observations. Most ‘global governance’ apologists who applaud the increasing use of benchmarking in policy research rarely acknowledge that the production of most indicators (like ‘rule of law’ indexes) is based upon fragile methodological foundations, and that the process of turning measurements into policy recommendations most often turns uncertainties and approximations into certainties... (Davis et al. 2012).

Indeed, it is precisely the construction of the doxa of a governable, manageable world that paradoxically the mono-disciplinarity of economics has resulted in: in such a world, actors that participate in its making have

to be selective and actively ignore inconvenient data, or, as the empirical example above illustrated, systematically disregard the development of some measurement tools versus others. As recently one of the METRO interviewees emphatically suggested, ‘it is art, not science’. This art of assembling knowledge, while actively and strategically constructing non-knowledge, is necessary in order to leave the epistemic authority of the solutions uncompromised (no matter how ‘clumsy’ these may be), as well as having the door always open to the construction of new problems and solutions once the previous ones fail.

Although the field of ignorance studies has put a lot of emphasis on classifying kinds of non-knowledge, it has so far not achieved a coherent set of ideas about *how* to investigate the process of producing non-knowledge. The most notable exception to this is Scheel and Ustek-Spilda’s (2019) work; the latter use the notion of enactment from STS, while also making references to the concept of controversies, and in particular the examination of cases of non-transfer of knowledge—the moments of distortion, reinterpretation and loss that may occur when ‘data move between people, substates, organizations, or machines’ (Edwards et al., 2011, p. 669). The attention to the particularities, representations and often visualisations (through graphs, maps and other visuals) that the enactment agenda allows could be seen as a helpful way of investigating the tools and effects of the production of ignorance.

To conclude, this chapter mobilised relevant literature and used empirical examples in order to offer two propositions: first, that instead of disciplinarity, global education governance is primarily dependent on a monodisciplinary knowledge production orientation; and second, that an investigation of metrological realism needs to focus on the social construction of non-knowledge as a vital component of studying the epistemic authority of transnational institutions. Perhaps a sceptical turn in the study of transnational regulation, evaluation and monitoring must lead to an ‘un-settling’ of the classic studies of the political use of statistical knowledge, and offer the promise of a more creative, at times even inconvenient, analysis of the unaccounted and thus invisible processes of the construction of non-knowledge that the making of quantification requires.

REFERENCES

- Anderson, K. (2014). *The learning metrics task force 2.0: Taking the global dialogues on measuring learning to the country level*. Retrieved on 1 September 2021 from <https://neqmap.bangkok.unesco.org/wp-content/uploads/2019/06/The-Learning-Metrics-Task-Force-2.0.pdf>
- Aradau, C. (2017). Assembling (non)knowledge: Security, law, and surveillance in a digital world'. *International Political Sociology*, 11(4), 327–342.
- 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, 2. <https://doi.org/10.1108/S0733-558X20210000074027>
- Bandola-Gill, J. (2020, November). The legitimacy of experts in policy: Navigating technocratic and political accountability in the case of global poverty governance. *Evidence and Policy: A Journal of Research, Debate and Practice*, 17(4), 615–633.
- 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.
- Beck, U., & Wehling, P. (2012). The politics of non-knowing: An emerging area of social and political conflict in reflexive modernity. In P. Baert & F. D. Rubio (Eds), *The politics of knowledge*. Routledge.
- Benavot, A., & Smith, W. (2020). Global learning metrics as a ready-made solution to a manufactured crisis. In A. Wulff (Ed.), *Grading goal four: Tensions, threats and opportunities in the SDG on quality education*. Brill.
- Bowker, G. (1995). Second nature once removed: Time, space and representations. *Time and Society*, 4(1), 47–66.
- Calkins, S., & Rottenburg, R. (2017). Evidence, infrastructure and worth. In P. Harvey, C. B. Jensen, & A. Morita (Eds), *Infrastructures and social complexity*. Routledge.
- 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.
- Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., & Borgman, C. L. (2011). Science friction: Data, metadata, and collaboration. *Social Studies of Science*, 41(5), 667–690.
- Fairclough. (1995). *Critical discourse analysis*. London: Longman

- 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 goals in the time of the SDGs* (PhD thesis). Facultat de Ciències Polítiques I Sociologia, Universitat Autònoma de Barcelona.
- Fourcade, M., Ollion, E., & Algan, Y. (2015). The superiority of economists. *Journal of Economic Perspectives*, 29(1), 89–114.
- Fukuda-Parr, S., & McNeill, D. 2019. Knowledge and politics in setting and measuring the SDGs: Introduction to special issue. *Global Policy*, 10(Suppl. 1), 5–15.
- 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.
- Grek, S. (2013). Expert moves: International comparative testing and the rise of expertocracy. *Journal of Education Policy*. <https://doi.org/10.1080/02680939.2012.758825>
- Grek, S. (2020). Prophets, saviours and saints: Symbolic governance and the rise of a transnational metrological field. *International Review of Education*, 66, 139–166. <https://doi.org/10.1007/s11159-020-09844-z>
- Gross, M. (2012). ‘Objective culture’ and the development of nonknowledge: Georg Simmel and the reverse side of knowing. *Cultural Sociology*, 6(4), 422–437.
- Gross, M. (2016). Risk as zombie category: Ulrich Beck’s unfinished project of the ‘non-knowledge’ Society. *Security Dialogue*, 47(5), 386–402.
- Gross, M., & McGoey, L. (2015). Introduction. In M. Gross & L. McGoey (Eds.), *Routledge international handbook of ignorance studies*. Routledge.
- Gunter, H., Grimaldi, E., Hall, D., & Serpieri, R. (2016). *New public management and the reform of education: European lessons for policy and practice*. Routledge.
- Guy Peters, B. (2017). What is so wicked about wicked problems? A conceptual analysis and a research program. *Policy and Society*, 36(3), 385–396. <https://doi.org/10.1080/14494035.2017.1361633>
- Hanushek, E., & Kimko, D. (2000). Schooling, labor force quality and the growth of nations. *American Economic Review*, 95(5), 1184–1208.
- Harvey, P., Jensen, C. B., & Morita, A. (2017). *Infrastructures and social complexity*. Routledge.

- IEG-WB. (2006). *From schooling access to learning outcomes. An unfinished agenda. An evaluation of World Bank support to primary education*. The World Bank.
- King, K. (2017). Lost in translation? The challenge of translating the global education goal and targets into global indicators. *Compare: A Journal of Comparative and International Education*, 47(6), 801–817. <https://doi.org/10.1080/03057925.2017.1339263>
- Klees, S. J., Stromquist, N., Samoff, J., & Vally, S. (2019). The world development report on education: A critical analysis. *Development and Change*, 00, 1–18. <https://doi.org/10.1111/dech.12483>
- Kleinman, D. L., & Suryanarayanan, S. (2013). Dying bees and the social production of ignorance. *Science, Technology & Human Values*, 38(4), 492–517.
- Larkin, B. (2013). The poetics and politics of infrastructure. *Annual Review of Anthropology*, 42, 327–343.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.
- Lawn, M., & Grek, S. (2012). *Europeanising education: Governing a new policy space*. Symposium.
- Mallard, G., & McGoey, L. (2018). Strategic ignorance and global governance: An ecumenical approach to epistemologies of global power. *The British Journal of Sociology*, 69(4), 883–909.
- McGoey, L. (2012a). Strategic unknowns: Towards a sociology of ignorance. *Economy and Society*, 41(1), 1–16.
- McGoey, L. (2012b). The logic of strategic ignorance. *The British Journal of Sociology*, 63(6), 553–576.
- McGrath, S., & Nolan, A. (2016). SDG4 and the child’s right to education. *NORRAG News*, 54, 122–123.
- Mundy, K., Green, A., Lingard, B., & Verger, A. (2016). *Handbook of global education policy*. Wiley-Blackwell.
- Naidoo, J. (2016). *Mainstreaming SDG4- Education 2030 in sector-wide policy and planning*, Paris: United Nations
- 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* (IRPA Working Paper GAL Series No. 5/2012).
- 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(5), 663–681.
- Smith, A. (2009). *Studying the government of the EU: The promise of political sociology*. Europa Institute Seminar Series. University of Edinburgh.

- Smith, W. (2019). 'One Indicator to Rule them All: How SDG 4.1.1 Dominates the Conversation and What it means for the Most Marginalised', *Annual Review of Comparative and International Education*, 37, 27-34
- Smithson, M. (2008). Social theories of ignorance. In R. Proctor & L. Schiebinger (Eds.), *Agnotology: The making and unmaking of ignorance*. Stanford 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.
- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 377–391.
- Stevens, A. (2011). Telling policy stories: An ethnographic study of the use of evidence in policy-making in the UK. *Journal of Social Policy*, 40(2), 237–255. <https://doi.org/10.1017/S0047279410000723>
- Tichenor, M., Merry, S., Grek, S., Bandola-Gill, J. (2022). Global public policy in a quantified world: Sustainable Development Goals as epistemic infrastructures. *Policy and Society* 41(4), 431–444.
- 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/CUE. (2013). *Toward universal learning. What every child should learn*. UIS/Center for Universal Education at Brookings.
- UNESCO. (2017). *Unpacking Sustainable Development Goal 4/Education 2030*. UNESCO.
- UNESCO. (2016). *First meeting of the SDG-education 2030 Steering Committee, 25–26 May 2016 (ED-2016/ED2030/ME/1)*. UNESCO.
- United Nations. (2015). *Transforming our world: The 2030 agenda or sustainable development*. UN.
- Verger, A., Parcerisa, L., & Fontdevila, C. (2019). The growth and spread of large-scale assessments and test-based accountabilities: A political sociology of global education reforms. *Educational Review*, 71(1), 5–30. <https://doi.org/10.1080/00131911.2019.1522045>
- Winner, L. (1986). Do statistics have politics? In D. Mackenzie & J. Wajeman (Eds.), *The social shaping of technology* (pp. 28–40). Open University Press.
- Wodak, R., & Meyer, M. (2001). *Methods of Critical Discourse Analysis*. Sage.

Wulff, A. (2018). SDG4 & education for sustainable development: Why we should celebrate an apparent failure? *World of Education*. Retrieved on 1 September 2020 from https://www.ci-ie.org/en/woe_homepage/woe_detail/16107/sdg4-education-for-sustainable-development-why-we-should-celebrate-an-apparent-failure-by-antonia-wulff

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Constructing Consensus by Data

1 INTRODUCTION

One of METRO's most startling findings was that international organisations are not merely 'centres of calculation', i.e. the organisations where numbers are produced. Instead, the majority of the experts we interviewed are former statisticians, whose main role is to act as brokers between different levels of governance and actors in the field. Brokerage in the global public policy arena involves primarily facilitating processes of socialisation, collective puzzling and creating interdependencies between a diversity of actors, be them experts, donors or national representatives in order to create conditions of collaboration, mutual trust and agreement. As I will show, alongside the work of measurement, this continuous brokerage work is equally key, if consensus on establishing global goals—and the routes to achieve them—is to be established.

Through IOs' bilateral relationships with participant nations and local communities, as well as with other IOs, research agencies, donors and funders, brokerage facilitates the adoption of different identities for different audiences and thus creates conditions of trust for achieving consensus. Second—and perhaps more importantly—brokering work creates the necessary bridge between technocratic and political accountability, essential in global monitoring programmes, giving large IOs further legitimation and symbolic capital as purveyors of technocracy *and* democracy.

Therefore, the focus of this chapter will be on the role of actors' socialisation, processes of collective puzzling, interdependency and brokerage in the production of goal-setting in education, in order to achieve consensus. After a brief overview of key literature analysing expertise and its main functions and effects, the chapter will move on to the brief examination of two empirical cases of expert brokering work in the field of global education governance: the first will focus on actors' socialisation by the OECD and its country reviews of education. The second case will explore UN's participatory turn, paying attention specifically to the work of UNESCO as a trusted education broker through the organisation of technical groupings and meetings for the purposes of measuring the education SDG.

2 THE OECD COUNTRY REVIEWS OF EDUCATION: THE CASE OF SWEDEN

Adopting a perspective that builds on sociological institutionalism (Lowndes, 2010), IOs are understood as purposive actors who, 'armed with a notion of progress, an idea of how to create a better life, and some understanding of the conversion process', have become the 'missionaries of our time' (Barnett & Finnemore, 1999; 712). However, this does not in itself explain what has transformed the OECD to one of the most powerful agents of transnational education governance. Martens (2007) has contributed to this discussion suggesting that the 'comparative turn'—'a scientific approach to political decision making' (2007, p. 42)—was the main driver of the OECD success in education governance globally. Through its education statistics, reports and studies, the OECD has achieved a brand which most regard indisputable. Despite a number of critical voices in the field (Brown et al., 2007; Prais, 2003), OECD's recommendations are accepted as valid by politicians and scholars alike, 'without the author seeing any need beyond the label "OECD" to justify the authoritative character of the knowledge contained therein' (Porter & Webb, 2004).

However, despite this context of increasing and deepening influence of its quantitative measures in education, the OECD's 'Reviews of National Policies of Education' show that the assumption that the OECD's influence is a mere result of its ability to decontextualize and compare is only but half the truth. Although the significance of the technisation of many—previously political—arguments in education cannot be

disputed, here I focus on a less discussed, yet important factor or the OECD success: this is the *sustained socialisation of policy actors within national contexts* through processes of policy translation and contextual adaptation (Checkel, 2005). As suggested by Checkel (2005), processes of socialisation entail intensive communication, regular meetings, as well as the emergence of mutual trust and shared commitment between actors who are involved in the ‘common project’. Socialisation leads to the construction of a common *esprit de corps*, defined as the acceptance and internalisation of new norms: ‘the right thing to do’. This, of course, is not always an orderly, observable process. Instead, it is a gradual, multi-layered process that is predominantly governed by a *logic of appropriateness*, meaning the adoption of institutional rules and norms that ‘regulate the use of authority and power and provide actors with resources, legitimacy, standards of evaluation, perceptions, identities and a sense of meaning’ (Olsen, 1998; 96). As I will show, what we observe in many countries around the world is the making of an almost absolute and indisputable consensus on the role and significance of the OECD as key in reshaping the academic, policy and public debate. Observing and evidencing processes of international and national actors’ socialisation as they take part in these institutional processes is considered an important intellectual tool in making sense of these new realities.

Second, in an attempt to illuminate *how* socialisation happens, Hugh Heclo’s notion of *collective puzzlement* (1974), as well as Clarke et al.’s (2015) conceptualisation of how policy moves, are both useful analytical tools. Both sets of ideas help to show how and why it is the coming together of various national and international actors that sustains and reinforces the numbers game, rather than solely the validity or strength of the numbers themselves. Over time (and the allowance of time is crucial here), at least in the field of education, international comparative assessments have created two crucial governing constructs: first, a common *language* using which diverse actors from the local, national and international ‘levels’ can communicate; second, a new governance system which in effect can be understood as ‘an incremental process reorienting the direction and shape of politics to the degree that (global) political and economic dynamics become part of the organisational logic of national politics and policy-making’ (Ladrech, 1994, p. 69). However, rather than top-down, this is a mutually reinforcing process; Sweden, the case in point in this section, was a key nation in establishing the work of international actors in education, like the IEA and the OECD, and is very active in

relation to European governance in education more generally (Grek & Lindgren, 2015).

In order to pre-empt critique, I do not claim that numbers are not important, or that their spectacle through naming and shaming (Nóvoa & Yariv-Mashal, 2003; Simola, 2005; Carvalho, 2012) is not an indispensable part of OECD's success. Nonetheless, the spectacle has a temporal dimension; it surprises and shocks. Thus, spectacles quickly come and go (think of the embargoed results for example, and the media attention the Programme of International Student Assessment {PISA} receives). Nonetheless, what follows the announcement of the results requires steadfast, diligent and zealous face-to-face policy work in order to carry the numbers deeper into the national imaginary and entrench them into the system. The OECD sustains and builds its policy work through the continuous crafting of its relationship with key education actors in other international organisations (Grek, 2014) and within national contexts.

But how do such processes of socialisation take place? One way to understand and analyse them is through the prism of policy learning defined as an 'updating of beliefs':

In public policy, we are eminently concerned with beliefs about policies ... This process of updating beliefs can be the result of social interaction, appraisals of one's experience (often of failure) or evidence-based analysis – or most likely a mix of the three. (Dunlop & Radaelli, 2013, p. 600)

Policy learning theory is certainly not new—from the seminal work of Dolowitz and Marsh on policy transfer (1996) and Haas' work on epistemic communities (1992), to the advocacy coalition framework (Sabatier & Jenkins-Smith, 1993) and the examination of the EU as a learning organisation (Zito & Schout, 2009), the literature on policy learning is large. Here I go a bit further back in time and focus on Hugh Heclo and his writings about governing as *collective puzzling*:

Politics finds its sources not only in power but also in uncertainty... Governments not only 'power'... they also puzzle. Policy making is a form of collective puzzlement on society's behalf; it entails both deciding and not knowing.... (Heclo, 1974, pp. 305–306)

According to Heclo, more so than politicians, it is the work of civil servants that is crucial in the making of policy; they are bestowed a

permanency that politicians do not have, in addition to experience and institutional memory, since ‘to officials has fallen the task of gathering, coding, storing and interpreting policy experience’ (Hecló, 1974, p. 303). However, policy work usually happens through interaction; according to him, ‘it is in interaction (that) these individuals acquire and produce changed patterns of collective action’ (Hecló, 1974, p. 306).

More recently, Clarke and colleagues suggest that policy is never a finished product, to be observed and transferred in a linear manner (2015). Instead, they suggest that,

‘When policy moves, it is always *translated*: that is, it is made to mean something in its new context. Policy is never a singular entity: it is put together – or assembled – from a variety of elements that are always in the process of being re-assembled in new, often surprising ways. (Clarke et al., 2015, p. 1)¹

Following both Checkel (2005) and Clarke et al. (2015), the OECD education policy work of the last 20 years has achieved a paradigmatic shift in the thinking and framing of education not only thanks to the cold rationality of numbers, but also crucially through the interpretation and adaptation of its recommendations in myriad venues and opportunities where local, national and international actors interact. Freeman sums up beautifully the impact of such iterative processes of collective learning:

This implies that learning is not simply an interpretative act, a process of registering and taking account of the world; it is, in a fundamental way, about creating the world. It is an active process of *making* sense (Weick, 1995). Similarly, just as we shop in order to discover what we want (and we might think of some kinds of political learning as “policy shopping”), so we read in order to discover what we think, not just what any given author thinks (Brown & Duguid, 2000). What emerges is *a conception of learning as an act of imagination, invention and persuasion as much as (or*

¹ The concepts of translation and assemblage have a strong footing in STS and especially Actor-Network theory. Åm (2016) criticised Clarke et al. for their use of the concepts without referring explicitly to STS. Indeed, an institutional approach does not marry very well with STS’s use of the concept of translation; although it would have been an interesting discussion, it is not possible to achieve this here. Therefore, the paper uses the looser term ‘interpretation’ to evoke the change of meaning and adaptation Clarke et al. (2015) persuasively discuss.

as well as) comprehension, deduction and assimilation. (2008, p. 15, my emphasis)

The next section empirically analyses the case of Sweden and the OECD, in order to show how such processes of socialisation and collective puzzling happen.

2.1 Socialisation and Learning in Governing: The OECD Reviews of National Policies for Education

As the OECD itself suggests, the ‘Reviews of National Policies for Education’ are one type of their range of activities that lead to analyses of education policy development and implementation (OECD, 2016). According to the OECD, there is involvement of Ministries as well as professional groups, researchers and others, in formulating and carrying out the work and in discussing the findings of the OECD review expert group that visits the country; thus, the circle of participating actors is wide and includes both national and international actors (OECD, 2016). The aim of the Reviews is ‘to improve the understanding of issues, implications for education policies and experience with the range of national policy options and strategies’ (OECD, 2016). Recent ‘National Policies’ include a high number of reviews from a diversity of countries; for example, the Netherlands, Latvia, South Africa, Dominican Republic, Russia, Scotland, Bulgaria, Korea, Ireland, Italy, Estonia, Lithuania, Kazakhstan, Chile and many others. Indeed, going back into the OECD archives, it is difficult to identify countries that have *not* had an OECD review of their education system.

Education policy reviews proceed in several stages: initially, there is preparation and completion of a background report by the country undergoing review, followed by a two-week mission by an external team of reviewers. The external team then prepares and completes the review report. This is presented at a 1 to 1½ day review session at the OECD Education Committee, when the Minister (with input from senior staff) comments on recommendations and conclusions of the review team and responds to questions of other countries’ delegates to the Education Committee (OECD, 2016).

The report of the external review team, edited to take into account the main points raised in the review session, is then published. According to the OECD, their scope is usually very broad with the goal to

provide recommendations on ‘effective policy design and implementation’. Generally the analysis covers ‘strengths and weaknesses which are primarily based on OECD’s collected data (from studies such as PISA, or earlier OECD reviews), national research, review visits to the country and OECD’s extended knowledge base’ (OECD, 2016). Finally, the programme of reviews consists of a follow-up. After a period of about two years, ‘authorities of the country concerned submit a short note to the Education Committee in which they report on progress and developments. Discussion takes place as a regular item in the agenda at a bi-annual meeting of the Education Committee’ (OECD, 2016).

2.2 The OECD Country Review of Sweden (2015) and the Foundation of the Swedish School Commission

The Swedish OECD country review of 2015 was not the first one in the country; another one had preceded it in 2011 (Nusche et al., 2011). However, in light of the negative PISA 2012 results, as well as the general downward spiral of Swedish education performance, it quickly led the Ministry of Education and Research (MoER) to commission the OECD for yet another report of the country’s education system. The objectives of the review were to

- 1) identify the main reasons for the decreasing trends in Swedish students’ performance; 2) draw on lessons from PISA and other benchmarking countries/regions with an expert analysis of key aspects of education policy in Sweden; and 3) highlight areas of policy and its implementation which might add further value to Sweden’s efforts to improve student performance. (OECD, 2015, p. 13)

The process followed the usual pattern: a background report prepared by the Swedish government, an OECD pre-visit which defined the key areas for review, an OECD team review visit to Sweden in October 2014, as well as a series of other exchanges with experts and stakeholders in Sweden and internationally (OECD, 2015). The two external experts in the team were Richard Elmore, Gregory R. Anrig Research Professor of Educational Leadership, Harvard Graduate School of Education, and Professor Graham Donaldson, the former Scottish HMI Chief

Inspector and the then president of the Standing International Conference of Inspectors (SICI)—Donaldson was one of the chief architects of the self-evaluation model in Scotland.

The OECD visit took place between 13 and 22 October 2014 and involved a number of meetings with key actors such as the Ministry; the Swedish National Agency for Education (Skolverket); the Swedish Schools Inspectorate (Skolinspektionen); the two teacher unions (Läraryrskörförbundet and Lärarnas Riksförbundet); academics in education research and teacher education (Stockholms universitet and others); the Swedish Association of Local Authorities and Regions (*Sveriges Kommuner och Landsting*); and visits in different municipalities and local schools (OECD, 2015).

The report uses quite damning language to describe the state of Swedish education: ‘no other country ...saw a steeper decline’ (ibid.; 7); ‘a school system in need of urgent change’ (ibid.; 11); ‘a position significantly below the average’ (ibid.; 27). On the basis of this discursive analysis of the text—which used a language that described a system in crisis—I interviewed key actors that contributed to the report. Their reflections on the process of how the Review was commissioned and its effects were enlightening.

Although they do not themselves use the term socialisation, all interviewees in their interpretation of the influence of PISA in Sweden offered a similar story of staggered events that followed one another; of the involvement of an ever wider set of actors; of the importance of the OECD experts in offering suggestions; and of the central role of the establishment of the Swedish School Commission as a forum of meeting, debate and learning for all the actors involved. Indeed, the title of report of the Commission, ‘Samling för Skolan’ (Gustafsson et al., 2017), denotes precisely the notion of ‘congregation’ or ‘gathering’—the meeting and consensus of different actors around the core of the commission’s study, which was the OECD numbers themselves. Numbers and data are central in the interviewees’ narratives, but so are the meetings, the debates and the continuous coming together of actors in socialising and learning events.

Interestingly, perhaps simultaneously with the rise of the OECD as the ultimate go-to expert organisation, we observe the slow decline of Swedish education research as valid and trustworthy enough to even

take part in the PISA data collection process—instead, Andreas Schleicher acquired an almost divine quality that matches closely the religious adherence to PISA in Sweden:

What has happened, you can go back to 2003 , TIMMS and PISA were at MidSweden university, now they are all run by the educational board (Skolverket). And they contract fewer and fewer education researchers for very little time to do some coding, to offer some comments. We were really independent from the government and at the time we were in a lot of the OECD meetings, we were involved. But now it is the educational board which does all that – and they don't have any researchers, they have project managers but they do not have researchers, they have government bureaucrats....But when Andreas Schleicher is in Sweden it is like we have a visit from God, it is very strange. (Academic 2)

As a result, education researchers do not have an alternative voice in Sweden anymore—when they take part, the majority of them is to validate rather than dispute the PISA results:

Today no one can [criticise PISA] really. PISA has in some sense got so much status that I don't meet many who can say we can contrast PISA – but a lot of people say we need to discuss the implications of PISA. (Academic 2)

Although the academic community appears to have lost its central position in informing policy, there appears to be a much more diverse and horizontal participation of different actors in policymaking, even if it involves a lot of 'cherrypicking'. Here, speaking about how the OECD report was commissioned, an interviewee, who later became central to its analysis, suggests:

Many Swedish organisations and persons, researchers, people in the professions [were asked to participate] and were listened to -very selectively of course- and did a lot of cherry picking of what they liked to hear and what they didn't like to hear – this is what politicians do. (Commission member 4)

What is important here are two developments that seemed to have dominated the Swedish education policyscape since 2000; the first one

was the unequivocal rise of the OECD as the golden standard of education research in the country (with the simultaneous downgrading of national education researchers); and second, the rise and broadening up of a debate about a system that was portrayed as in crisis. This picture, given the history of Sweden as a model European education system throughout the twentieth century, in addition to the success of close neighbours, such as Finland, became symbolic of a marked shift in the need to socialise and ‘educate’ all relevant actors about the critical need for change. That process began slowly since the mid-2000s, but became cataclysmic after the damning PISA 2012 report. Indeed, it was PISA 2012 that became the primary reason for launching the Swedish School Commission:

It was a response to the OECD report. If I can give you a bit of the timeline : in December 2013 we have the PISA report, week after that there was a big debate at the parliament about the school crisis. There after Björklund invites the OECD to write the report, even before the report is released and they organised this school commission with Anna Ekström - now the chair is Jan –Eric Gustaffson. Their task was to study the report of the OECD in order to make a Swedish analysis, do we agree what is the to-do list, but this commission has been criticised as being only in favour of this particular view that the PISA results are the only ones that show the truth about Swedish schools today. (Academic 1)

Indeed, the task of the Commission was set out as follows: ‘partly based on the OECD’s recommendations, the schools commission will submit proposals aimed at improving learning outcomes, teaching and equity in Swedish schools’² (Swedish Government, 2015). Indeed, the OECD and its recommendations were central to this debate and in many ways, framed it; this then instigated the work of the Commission that was purposefully staffed by a broad range of actors and that met regularly over two years in a process of learning, socialisation and translation of the OECD recommendations to national policy.

The Swedish School Commission met regularly for two years. Its members were asked to look at evidence and draw conclusions about the direction of travel for Swedish education. Interviewees described these meetings as learning opportunities for all participants involved. They

² For a detailed list of its members see here: <https://pasisahlberg.com/news/swedish-school-commission/>

described the Commission as broadly reflecting the wider public and policy debate in Sweden and suggested that its priority is to take the time necessary to offer a ‘Swedish solution’, nonetheless following closely the OECD research and recommendations. Again, in their narratives, they never claim that the OECD data are not central; on the contrary, they describe OECD data as the ‘spine’ that holds them all together. However, they also suggested that there was a national ‘filtering’ process that took place through their meetings, and that was necessary for the interpretation, adaptation, persuasion and at the end adoption of the OECD perspective.

To conclude this section, the OECD Swedish country review of 2015 and the set-up of the Commission that followed represent an illuminating case of the kind of processes of socialisation of actors that was discussed earlier on: in this case, the OECD was invited to enter a national system and combine its quantitative knowledge with a more qualitative perspective, gained from a two-week fieldwork visit, discussions with local actors, as well as a detailed background report supplied by the government of the time. If actors’ socialisation and collective puzzling via the spectacle of country rankings was the state of affairs prior to the rise of a field of global public policy with the SDGs, the UN’s participatory turn, where we will next turn, heralded a whole new era on the role of experts as brokers and particularly UNESCO’s influence, as will be charted in the chapter’s following sections.

3 UN’S ‘PARTICIPATORY TURN’: QUANTIFYING WHILE DEMOCRATISING

As 2015 was approaching, it became increasingly clear that the MDGs would not be achieved (Fukuda-Parr, 2017). One of the main causes of this failure was seen as the top-down UN structure that governed the goals. In addition to the failure to establish an effective governing architecture, the calls for decolonising and democratising the global public policy arena were also multiplying at the time and gaining increasing momentum. Hence, apart from establishing ambitious goals in themselves, another key ambition was to alter their governing architecture (compared to how the MDGs were organised), by democratising it and allowing countries to have a much stronger say:

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)

What is vividly illustrated here are two key tensions embedded in setting up the new global monitoring system: on the one hand, there is a clear break with a top-down, global North-centric view of sustainable development and the promotion of more equal and democratic participation by the countries who would be most affected by these systems. On the other hand, proclaiming such an inclusive design in setting up the monitoring system was seen as risking technical challenges and undermining the authority of expertise, as it would necessarily need to involve countries in the politics of measurement in a much more direct way.

What is of interest is that this 'participatory turn' of the UN monitoring system did not merely occur at the level of procedural backstage politics but rather, it was embedded in the key document establishing the SDGs. The flagship document of the Rio Conference—*The Future We Want* was at its core a political declaration of inclusivity of the different voices into the governance through but also of the Sustainable Development Goals. 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 policy-makers but also a range of other stakeholders (such as the civil society and national representatives) was seen as necessary for the success of the Sustainable Development Goals. *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 were 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)

Therefore, the SDGs journey was proclaimed as 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 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 in order to achieve consensus—the latter being the key underpinning principle of the new framework. Taken together, these two documents clearly show how the SDGs changed fundamentally the role and practice of expertise: what we observe here is that, alongside the need for technical knowledge, experts are required to work closely with country representatives and a range of other actors in order to achieve agreement on the goals. This extends the kinds of qualities that expert work involves: apart from needing to be statistically and technically highly competent, IO experts would need to also be successful in persuading other actors and securing support and buy-in, so as to be allowed to push on with their work. Surprisingly, what we found in the METRO project was that most of the experts involved in these processes were former statisticians (therefore, they had the technical capacity to understand the issues involved), however, the vast majority of the work that was required of them was to foster these relationships and broker agreement between a very wide diversity of actors with diverse ideas and interests.

These are some of the reasons that the introduction of this ‘participatory’ approach to statistics was not straightforward. The production of globally comparable statistics is a very complex and demanding technical process that cannot always be adjusted and made to fit with actors’ disagreements and political persuasions. One way in which these tensions were resolved—at least rhetorically—was the UN’s devising of 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. Thus, 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 were 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 fact, the technical group responsible for the indicator development—the IAEG-SDGs—set up ‘country ownership’ as one of their key goals. Hence, it is clear how this highly technical body is also required to act as a broker of relationships and consensus-making with participant countries, rather than merely do statistical work:

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 appear to resolve the technocratic and the democratic tensions of the SDGs. On the other, focusing on ‘ownership’ does not completely surrender quality standards of the indicator development, but communicates the need for countries to negotiate measurement as both a technical process and a political process of deciding on policy prioritisation. Thus, experts are required to maintain sufficient levels of technocratic accountability to reap the benefits of the ‘epistemic virtue’ (Daston & Galiston, 2007) of numbers (such as standardisation, objectivity and universality), while combining their technical capital with navigating important political calculations, such as securing consensus-building, promoting collective action *and* ensuring the political acceptability of the monitoring process and all the decisions needing to be taken therein. As I will discuss in

the next section, this is the kind of work that the UNESCO Institute of Statistics did, as it brought together not only its epistemic authority in the field but also its reputation as a trusted IO in the eyes of countries of the Global South in particular.

4 EXPERTS OR BROKERS? UNESCO INSTITUTE OF STATISTICS AS A TRUSTED ACTOR

As discussed in the previous section, the SDGs captured the imagination of a wide set of actors in the field, since they purposefully allowed multiple ‘entry points’ in their world: on the one hand, they emphasised the use of technocratic and management principles to create an objectified and measurable field, while also proclaiming to be bottom-up, grass-roots and transformative, distinct from older Western-liberal ideas and practices (Waldmüller et al., 2019). Such an open framing of the SDGs allowed them to move and adapt much faster than previous monitoring exercises, no doubt partly due to the malleability and flexibility of the monitoring framework itself. Thus, the scope and complexity of the SDGs lend itself to a focus on the structures and interlinkages between data, actors and politics, precisely the processes evidenced through the making of the SDG4.

Thus, this section turns the spotlight onto these key people, who, although occupying technical positions, have also been given a strong mandate towards achieving consensus among the wide diversity of participants in what is often referred to as the ‘indicator debate’. Here I focus on the struggles, tensions, as well as the transformations of actors’ epistemic, highly technical knowledge capital into a set of practices that focused primarily on brokerage and on achieving consensus. Crucially, such brokerage practices do not lessen the significance of quantification and specifically the ‘indicator debate’; instead, they promote the monitoring agenda and, through their unfolding, co-opt a wide range of actors.

Thus, in order to demonstrate the workings of brokerage and consensus-making in practice, it is useful to discuss the work of two key indicator groupings, namely the Technical Cooperation Group (TCG) and the Global Alliance for Monitoring Learning (GAML). While this section derives some data from the extensive online documentation of their regular meetings, it is primarily based on actors’ own voices as

participants of these meetings and as key active participants in these groups.

Although 2015, the year the SDGs were launched, seemed like the dawn of a new era for the global education community, it left a number of issues open—among them the so-called *indicator debate*. The Education 2030 (Incheon Declaration, 2015) had established four levels of indicators (global, thematic, regional and national). The first level included up to 11 global indicators, negotiated in a series of meetings of the Inter-Agency Expert Group on SDG Indicators (IAEG-SDGs).³ In light of the unequal development of these indicators,⁴ and often the unavailability of data and lack of coverage on a global scale, IEAG-SDGs implemented a 3-tier classification tool,⁵ which categorised indicators depending on their robustness according to internationally established methodologies and standards and the regularity of data production at country level. Importantly, IAEG-SDGs also identified a number of custodian agencies they deemed responsible for the development and refinement of such indicators. In the case of education (SDG 4), the UNESCO Institute of Statistics (UIS) became the responsible entity for 9 out of 11 indicators and was tasked to share the responsibility for the other 2 with UNICEF and the OECD. Given the initial classification of a number of metrics as tier 2 and tier 3 indicators (e.g. indicators for which data are not regularly produced by countries or for which measurement standards are not yet available), their refinement and production rapidly become a priority for UIS, who, as discussed below, perceived their organisational legitimacy and reputation as being closely tied to achieving both technical solutions

³ The Inter-Agency Expert Group on SDG Indicators (IAEG-SDGs) was created on 6 March 2015 by the United Nations Statistical Commission at its forty-sixth session. For more information, visit <https://unstats.un.org/sdgs/iaeg-sdgs/> [accessed 9 June 2020].

⁴ The development of indicators is unequal in terms of their being at very different levels of development, ranging from concrete ones (supported with data) to those described as *tier 3* indicators (see next footnote), i.e. lacking an established data infrastructure to measure them.

⁵ An indicator in *tier 1* is defined as being conceptually clear, having an internationally established methodology and available standards, with data being regularly produced by at least 50 per cent of countries in every region where the indicator is relevant. An indicator in *tier 2* is conceptually clear has an internationally established methodology and available standards, but data for this indicator are not regularly produced by countries. An indicator placed in *tier 3* is one for which no internationally established methodology or standards are yet available, but are being (or will be) developed or tested (UN 2020).

as well as the consensus of the participant actors about the pertinence and suitability of the indicators under consideration.

Given the complexity of the endeavour, but also in order to guarantee the participation of a wide range of stakeholders, two ad hoc mechanisms/working platforms were created with a view to advancing the development and production of SDG 4 global and thematic indicators. One was the ‘Technical Cooperation Group on the Indicators for SDG 4’ (TCG) and the other was the ‘Global Alliance to Monitor Learning’ (GAML).

The former was established in 2016, being conceived as a space for discussion as well as a technical platform to support UIS in the implementation of the thematic indicator framework. TCG is composed of regionally representative UNESCO Member States, as well as representatives of different IOs (UNESCO, UNICEF, OECD and the World Bank), civil society organisations and the co-chair of the Education 2030 Steering Committee.

GAML, on the other hand, was also created in 2016, being originally defined as an ‘umbrella initiative to monitor and track progress towards all learning-related Education 2030 targets’ (UIS, 2016, p. 49), and tasked with the development of tools, methodologies and shared standards to measure learning outcomes in the context of SDG 4. 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. Similar to TCG, GAML operates by definition in an open and participatory manner, with decisions being made through consensus.

Importantly, these platforms did not emerge in a vacuum. On the contrary, and as briefly discussed in the previous chapter, both of them were born out of already existing initiatives launched during the run up towards the approval of the Education 2030 Agenda. More specifically, TCG represents a continuation of the Technical Advisory Group (TAG) established in 2014, chaired by UNESCO and including experts from a range of education-related multi-lateral agencies. GAML, in turn, was a successor of the Learning Metrics Task Force (LMTF), launched in 2012 and envisaged as a multi-stakeholder partnership co-convened by the Center for Universal Education (CUE) at the Brookings Institution and UIS. For both TCG and GAML, there is extensive online documentation of their regular meetings, but they only really come alive in

the voices of individual participants of these meetings. According to two interviewees,

The [technical advisory] group that led the developments, that gave us the current indicators for SDG 4 was a precursor of TCG and GAML. And it was essentially the same composition ... that group basically just renamed itself as the TCG. They worked on the indicators that were then adopted. (UIS 1)⁶

The Learning Metrics Taskforce was another space where conversations were held and where I think most of the big actors in the global education policy space were somehow represented. I mean if you look at institutions involved ... so I think that also contributed to a consensus building that's become really hard to resist. (Civil society 1)

Nevertheless, despite a certain path dependency, the (re)formation of both TCG and GAML entailed a procedural shift *vis-à-vis* their own precursors. Both initiatives were explicitly set up in the understanding that they would be subject to a transparency mandate and were expected to operate in a democratic, equitable and inclusive manner. In this sense, both platforms and spaces are subject to a dual form of accountability—they are held responsible for the success of their technical work, but also judged in terms of the quality and inclusive character of their deliberations. To put it differently, both spaces are characterised by an inbuilt tension between the technical and the political accountability of the whole endeavour. An analysis of the work of the two indicator groups is therefore a productive space which serves as the canvas for mapping the struggles of the actors' positionings and efforts for producing consensus in the field. This tension is primarily evidenced through the centrifugal forces of technocracy, versus the perceived need for SDG4's inclusivity. As a number of my interviewees suggested, this has created a set of quality assurance problems that, although extant before, were never quite as prominent as in the case of the Education 2030 Agenda (SDG 4):

I think the way that UIS and GAML are trying to manage this is, say, if countries want to submit their national assessment data, that's fine, no

⁶ To protect interviewees' identities, I refer to them here merely giving an indication of their background (e.g. UIS, World Bank, OECD, UNESCO or Civil Society), with numbers merely serving to distinguish between multiple respondents from the same institution.

problem. But I think we're trying to manage by then saying, but we are going to put your data through a quality control process. And that's going to tell us technically how strong the data are. It doesn't mean we're not going to publish it, but the data may be published with an asterisk or footnote or in a slightly different way to signal to the viewer of the data that this is not exactly the same as say a TIMSS score or a PIRLS score⁷ ... it's let's get them in the door, let's just get it started, get people used to the habit of data on learning, then gradually raise the bar. (World Bank 1)

However, not all IOs share this tendency of prioritising inclusion over data robustness:

There is no clear-cut answer to this, I think it's a very difficult dilemma. But it also reveals a very different approach between UNESCO and OECD on how to respond to this. And we have encountered this not just in the outcomes metric but also in a way with general statistics. The UNESCO philosophy being we need to be open, we need to accept the constraints and at the end of the day it's better to have something than to have nothing. At the OECD I have taken a very different approach to this. For me, the most precious currency is trust. If I know that policymakers do not trust the data or don't trust them to be comparable, the whole thing is of very little value to me, because I want these things to be actually having an impact on this. So basically UNESCO and OECD use the same data source on the administrative side at least. With UNESCO the tables are full; for the OECD we have half of the cells filled with an M, which means actually these data aren't good enough. (OECD 1)

The latter quotation is exceptionally telling in terms of the work that numbers do in the construction of coalitions of actors, even when the data are not there at all. This is precisely the political function of numbers; even in their absence, they create the conditions for consensus and coalition-building. On the one hand, UNESCO appears to be using

⁷ The acronym TIMSS stands for Trends in International Mathematics and Science Study. PIRLS stands for Progress in International Reading Literacy Study. TIMSS and PIRLS are international assessments that monitor trends in student achievement in mathematics, science and reading. Both are conducted by the International Association for the Evaluation of Educational Achievement (IEA), an international cooperative of national research institutions, governmental research agencies, scholars and analysts. For more information, visit <https://www.iea.nl/studies/iea/> [accessed 7 May 2020].

statistical data as the means to mobilise an ever greater number of countries to participate in the global measurement operation. UNESCO'S primary 'currency' is inclusion and equal participation of all actors in the policy process; it uses numbers as a symbolic emblem of the belief of the organisation in more democratic and transparent processes of transnational education policymaking and monitoring. The OECD and the World Bank, on the other hand, still appear to be immovable from their core technocratic tenet of 'trust in numbers'; they use peer pressure to encourage countries to conform and participate. The symbolic use of the 'M' to denote missing data is not an empty cell; it symbolises in many ways the peer pressure and governing function that numbers have.

However, I do not wish to present these organisations as being in any way monolithic. The majority of actors I interviewed, even ones from the same organisation, often gave divergent views of their organisation's approach to data robustness and validity. However, it is precisely the contentious issue of the conflict between the technical and the political accountability of the monitoring tool which has been the breeding ground for the emergence of the 'metrological field' that governs transnational education. This field is inhabited by individual actors who assume different positions, sometimes following the culture of the organisation that employs them, but also—indeed often—not. These actors use their accumulated epistemic capital in order to transform it into a brokering device that facilitates their visibility, authority and legitimation in the field. More often than not, my interviewees cited their own career trajectories, values, frustrations or aspirations as the reasons which led them to take the position they had assumed; these positions are not permanent and solid. They often change in the face of developments in the 'field', i.e. the positionality, advancement and withdrawal of other actors involved in it. This conditions not only how they act, but also how they present themselves; style and substance can be of equal weight here.

4.1 *The Role of Meetings*

Observing the process and practice of these groups' gatherings is perhaps the most telling material evidence of how meetings contribute to the production of consensus around numbers and the policy directions that accompany them. 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 my

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. There is a clear-cut distinction of participants from the Global North, whose presence and contributions dominate the meetings, while 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. Thus, the space of the meetings becomes the visual manifestation of those who carry symbolic capital (and exercise authority) and those who do not, and whose *lack* of symbolic capital ironically enough also becomes a source of strategic positioning, since their agreement to the proposed agenda is required for the process to move on.

Further, the ambiguity and informality of the process, despite being an issue for some in the room, become a valuable, malleable tool in ensuring participation, while at the same time also pushing on with a specific, pre-determined agenda:

This was a big argument in the [removed for anonymity purposes] meeting two years ago. Because initially the [removed for anonymity purposes] ended up being a meeting of all the different actors in the assessment field, and they were kind of fighting among each other trying to frame their own assessment as the best for any SDG monitoring efforts. And this of course means that you have quite a lot of conflict of interest in the room. So one of the things that we as [removed for anonymity purposes] tried to say quite early on was that for this to be a well-functioning body that can actually do some work we would need to be quite clear on how decisions are made. And 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)

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. This kind of peer pressure, the discourse of crisis and the need for active involvement despite failings and malfunctions, trumps any hesitations about the process itself. In fact, as we see below, the process is informal enough to invite the complainant to try and sort out their own complaint:

[removed for anonymity purposes] then proposed that what if we had a strategic coordination committee that could approve the agendas in advance and that could try to ensure that this works well. And then [removed for anonymity purposes] was 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 [removed for anonymity purposes] 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. So we still feel like somehow we have to be in the room. (Civil society 1)

Chairing the meetings, although seemingly simply an administrative task, can also play an important role not only in how the meeting is run, but also in the conclusions that are drawn. Some of my interviewees suggested that the choice of chairpersons is strategic and aims to avoid (or prompt) specific actions, or the divergence of the problem to where the IOs want it to be:

Then the chair of the session will basically just wrap up, often without any reference to how things will move on. And this is enabled by the fact that [removed for anonymity purposes] is very seldom chairing the sessions herself, but she asks other people to chair. So you would have for instance the representative of the Australian Department for Trade and whatever it's called, foreign affairs and trade, I guess, chair a session. And then he will not in a way be expected to do the wrap-up in terms of follow-up, but he's really just brought in as the one who's facilitating the session. And then [removed for anonymity purposes] is doing a concluding statement of some sort, where she often has a PowerPoint and she would outline the next steps. But it's always completely unclear how the critical input from the group is really going to feed in or shape things. (Civil society 4)

Finally, as suggested earlier, the SDGs have prioritised consensus-making and the ‘democratisation’ of data as key in any forward-going process. Given the power asymmetries and the often informal management of the process, such efforts are frequently interpreted as a symbolic gesture that might even threaten the *doxa* of ‘trust in numbers’—they are, in a sense then, a ‘hetero-doxo’, a necessary deviation from accepted standards to sustain the ambivalence and multiplicity of the field.

It is precisely this conflict between methodological robustness and democratic participation that seems to set the SDG 4 wheels in motion. Although IOs seem to have the epistemic capital to drive the process, the need for participant countries to agree and approve suggests that other forms of capital are key, too. This leverage that participant nations and other civil society organisations have can be seen as problematic at times:

These are the professional standards of the measurement community. These are the instruments that are going to help you align your assessment with what good practice is, we’re recommending you use it. But it feels like we’ve gone into a whole second phase of, and now we’re going to get all the countries together, and we’re going to get everyone’s buy-in. And it keeps going back to this theme of democracy and voice. Is there such a thing as too much democracy, too much consultation? At what stage do you say we just have to run with this, this is what it is? (World Bank 1)

To conclude, one of the greatest difficulties of analysing expertise within the transnational metrological field is its complexity, dynamism and multiplicity. Although quantification has dominated global governance as the new unequivocal *doxa* of planning the future, it is precisely its open imbrication into political struggles that have transformed it into a powerful governing tool: the work of expert brokerage has been a key tool in achieving this balance and this transformational power.

5 DISCUSSION

The sociology of quantification has richly explained the ways that the work of counting is a deeply political process, despite its claims to rationality and objectivity (Merry, 2016). Indeed, quantification needs to ‘de-politicise’, in order to claim its legitimacy and authority; this is the main reason ‘why International Organisations hate politics’, according to the recent book by Louis and Maertens (2021). Indeed, there have been

plenty of detailed accounts of the processes of technicisation that social problems often go under, in order for experts to render them technical, and thus factual and neutral, and distinct from obstructive political struggles and ideologies (Wood & Flinders, 2014). Similarly, Diane Stone uses the term ‘scientization’ to describe the processes of transforming social issues into problems amenable to the scientific cause-effect relationship; the latter is seen as authoritative enough to control or even reduce uncertainty and risk (Broome et al., 2018; Stone, 2017). Of course, there is nothing a-political in such processes of rendering social problems as technical issues; on the contrary, technicisation is deeply political work that involves decisions about what to count and what to ignore, which variables to disaggregate and which not (some of these wilful acts of performing ignorance were already discussed in Chapter 3), and how much to spend on collecting and analysing information.

This chapter discussed the political work of technicisation as a project that does not ‘land’ into policy contexts as a top-down agenda, sent from some unknown ‘centre of calculation’, but one that is open to contestation and negotiation with participant countries. One of the underpinning assumptions of the epistemic power of quantification and its influence has been the separation of the spheres of science and politics (Lahn & Sundqvist, 2017). Yet, this positioning of measurement as objective and devoid of politics is increasingly challenged not only on the grounds of ethics and on democracy, effectiveness and efficiency, but also on its ability to win ‘hearts and minds’. 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. Thus, through the discussion of the empirical cases of the OECD’s education country reviews, the UN’s participatory turn, as well as the function and role of technical groupings and their meetings for the production of the education SDG, I showed how expert work is not merely the technical and statistically robust process that quantification promises, but it has come to deliver a function that politically is even more significant and necessary if global comparative metrics and their associated policy goals, are to be achieved: this is the work of socialisation, interdependence and brokerage that many of the experts METRO interviewed are asked to perform in order for the global goals machinery to move, one clog at a time.

First, through the case of the OECD's country reviews, I showed that, rather than simply offering what has been seen as fast policy solutions (Lewis & Hogan, 2019), the OECD painstakingly enters national sites and works with local actors to create conditions of belonging; that is, it creates conditions fruitful for collective puzzlement, socialisation and policy translation as Hecló (1974) and Clarke et al. (2015) suggest. The set-up of the Swedish School Commission with a remit to study the OECD report in detail and offer recommendations for reform could not have been a better example of quantification as an Ianus-faced process of both the simultaneous de- and re-politicisation of the problem of perceived under-performance in education. Although IOs are the usual suspects in the scholarship that focuses on the production of global comparative metrics, I showed how national actors were equally central in supporting, sustaining and even strengthening these processes. Indeed, some of the interviewees, even when critical of the OECD work, were ready to acknowledge that the OECD sparked a debate that would not have happened otherwise. Progressively, since the mid-2000s, the OECD became an undisputed expert organisation in Sweden, and indeed, as couple of interviewees suggested, a 'production force'. Close and sustained work with the Ministry, in combination with touching a nerve with the Swedish public (with quotes by Schleicher, such as 'Swedish schools having lost their soul'⁸) were key ingredients of this success. In the case of the OECD and Sweden then, ironically perhaps, 'governing at a distance' (Cooper, 1998) appears to require a strange sense of proximity: arguably, these conditions of actors' socialisation and policy translation are necessary for the kind of paradigmatic policy shift that quantification has led to in global public policy.

Through processes of collective puzzling and social learning, experts bring forward a new mode of regulation which draws on and supports the 'data dream' by providing it—at least in some systems—with what it lacked before: a sense of belonging and 'ownership' of the project. This is how international reform agendas enter national policy spaces and shape them through slow, continuous and consensual build-up of the new, common *esprit de corps*—the inescapable 'right thing to do' (Meyer, 2005). It is these processes of learning and socialisation that embed the international much deeper into the national consciousness, one

⁸ <https://www.theguardian.com/world/2015/may/04/sweden-school-choice-education-decline-oecd>.

often traumatised by the exposure that the damning global comparative data may bring.

The work of expertise, therefore, has undergone changes that may be seen as emblematic of a paradigm shift not only in the regulation of education, but also in regulation per se. Crucially, as I showed, socialisation and the learning that it produces does not merely entail the learning of facts. It is *constitutive*, generating or strengthening trust, commitments, identifications and loyalties—it embodies, as Hunter has fittingly described, ‘the connective tissue of governing itself’ (Hunter quoted in Newman, 2012).

Indeed, the UN’s participatory turn is evident of very similar tendencies and processes taking place, with the involvement of national actors, at global sites of measurement and decision-making. No matter how important the socialisation of national actors in these processes is, it has to be matched and strengthened by the expansion of connections and interdependencies grounded in the new governing paradigm of the global goals, as exemplified in the production and measurement of the SDGs. As the chapter discussed, this participatory logic was embedded in the SDGs from their inception with important consequences for the governance structures of this framework, as well as the implementation of the framework on the country level. Consequently, one of the key global IOs, the UN—and subsequently all the IOs that work with it—was driven not only by the technocratic logic of quantification, but also the demands for participatory governance.

From their inception, and in contrast to the MDGs and all other global monitoring programmes that preceded them, the SDGs were designed to be both a highly technocratic monitoring programme of ‘governing by numbers’ (Miller, 2001) *as well as* a participatory project aimed at assuring the participation of countries and communities. Such a double focus on both democracy and technocracy has been challenging for the experts within the IOs, as the technical decisions that had to be made were dependent on a concurrent process of their de-politicisation as technical matters that required data solutions, as well as re-politicised as issues that brought participants together in search of consensus. Such a balancing act required careful expert brokering work, as prioritising one over the other would risk the loss of momentum and support: for example, as the case of the MPI indicator showed, 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, technical considerations had to often be mobilised as tools of distraction, especially when difficult political decisions had to be made and no consensus was in sight.

This context, where the stability of objectivity was replaced by the fluidity of the continuous consultation processes, shaped to a large degree the new expert work of brokering. What METRO found was that numbers are no longer ‘fixed points’ (cf. Lahn & Sundqvist, 2017) but fluid entities that could always be improved, changed and mobilised in different ways. Experts had to mobilise not only their technical and epistemic capital but also a range of capitals at their disposal, such as the use of evocative language, beautiful data, marathon sessions and Global South participants flown around the world to attend—and thus legitimise—yet another meeting.

Finally, the case of the SDG4 is illustrative of what expert brokering work involves. For a start, although SDG 4 could be seen as a prime example of a transnational soft regulatory instrument (in the tradition of ‘soft’ law, i.e. best practices, expert standards, rankings, ratings, audits, quality assurance and the like), as I showed, it is also *substantially different* from other quantification exercises. The construction of SDG 4 represents a leap in the practice of transnational soft regulation because, although prescriptive, it also appears as transparent, pluralistic, open and developmental—consensus-making is prioritised by experts and data collection and validation processes are required to be ‘democratic’. I have described the ways in which the centrifugal forces of technical and political accountability have given shape to expert actors’ positionings and political work within it.

Nevertheless, these processes are not smooth and linear. As the empirical material shows, they involve 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. 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.

At the heart of this chapter are the paradoxes and the multiple ambivalences that quantification brings to transnational governance. 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. Indicator frameworks and all other subsets of numerical work create what Bourdieu (1977) terms a ‘linguistic market’. While some actors have the epistemic purchase to own and control most of this market, many others, as we have seen, are there—knowingly and willingly—to consume this lingua franca of numbers and transport it back home. Second, numbers’ underlying use as the new *doxa* of transnational governance legitimates a whole series of informal and ad hoc arrangements, all accepted and all approved in the name of the multiple global crises and the need to construct as broad a consensus as possible; in this fluid and dynamic arena, even bad quality data or even no data would do.

As we have seen above, the SDGs identified a specific failure of all previous statistical large-scale projects to bear fruit and developed a manifest governing programme to influence the behaviour of participating actors—and by ‘participating’ I do not mean only national ‘generalists’ but also highly technical elite experts who are now asked to expand their set of skills and adapt to this new governing reality. It may be that interventions still appear restricted to pushing (and largely financing) the statistical capacity for nations to produce data for governing; nevertheless, this step is seen as (and indeed is) key in achieving ‘transformative’ change. In terms of expert work, it creates different types of contributions (from the highly technical to the politically strategic and diplomatic) that require horizontal relationships between actors that are not fixed but are continuously negotiated and shared. Thus, expert work facilitates the emergence of a global public policy field that transcends the national/international/state/non-state divides.

To conclude, 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. In this metrological space or *field*, objective relations are structured by the distribution of economic, epistemic or cultural resources ‘which are or may become active, effective, like aces in a game of cards, in the competition for the appropriation of scarce goods of which this social universe is the site’ (Bourdieu, 1989, p. 17). The perceived weak positioning of Global South actors in the process is a telling example of this.

As a result, the non-existence of any ‘rules of the game’ in this field 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. Thus, quantification is key in the production of transnational governance, as it represents the unfolding development of ‘product and process’, constantly moving with that which it seeks to move. Instead of analysing expertise as solely a process of *depoliticising* social problems through the imposition of a measurement agenda, we observe a process of *re-politicisation* of policy problems by making them knowable and actionable through expert brokers. Expert work in this context represents an ‘act of performative magic’ (Bourdieu, 2000, p. 243) that is vital to the building of global public policy, as experts attend to and navigate the contested ideas and values which infuse the everyday realities inhabited by all participant actors. Expert work, therefore, as both ‘product and process’—despite all its contestations and failings—shape that which it classifies, generates particular scripts of action and reconfigures policy problems and issues in ways that invite certain possibilities for deliberation and allow the production of ‘consensus by data’.

REFERENCES

- Barnett, M. N., & Finnemore, M. (1999). The politics, power and pathologies of international organisations. *International Organization*, 53(4), 699–732.
- Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge University Press.
- Bourdieu, P. (1989). Social space and symbolic power. *Sociological Theory*, 7(1), 14–25.
- Bourdieu, P. (2000). *Pascalian meditations*. Stanford University Press.
- Broome, A., Homolar, A., & Kranke, M. (2018). Bad science: International organisations and the indirect power of global benchmarking. *European Journal of International Relations*, 24(3), 514–539.
- Brown, S., & Duguid, P. (2000). The social life of information, *Work Study*, 49(4). <https://doi.org/10.1108/ws.2000.07949dae.002>
- Brown, G., Micklewright, J., Schnepf, S. V., & Waldmann, R. (2007). International surveys of educational achievement: How robust are the findings? *Journal of the Royal Statistical Society Series A-Statistics in Society*, 170(3), 623–646.
- Carvalho, L. M. (2012). The fabrications and travels of a knowledge-policy instrument. *European Educational Research Journal*, 11(2), 172–188.
- Checkel, J. T. (2005). International institutions and socialisation in Europe: Introduction and frame work. *International Organisation*, 59(4), 801–826.

- Clarke, J., Bainton, D., Lendvai, N., & Stubbs, P. (2015). *Making policy move: Towards a politics of translation and assemblage*. Policy Press.
- Cooper, D. (1998). *Governing out of order: Space, law and the politics of belonging*. New York University Press.
- Daston, L., & Galison, P. (2007). *Objectivity*. Zone Books.
- Dolowitz, D., & Marsh, D. (1996). Who learns what from whom: A review of the policy transfer literature. *Political Studies*, 44(2), 343–357.
- Dunlop, C., & Radaelli, C. (2013). Systematising policy learning: From monoliths to dimensions. *Political Studies*, 61, 599–619.
- Freeman, R. (2008). Learning by meeting. *Critical Policy Studies*, 2(1), 1–24. <https://doi.org/10.1080/19460171.2008.9518529>
- Fukuda-Parr, S. (2017). *Millennium development goals: Ideas, interests and influence*. Taylor & Francis.
- Geertz, C. (1980). *Negara: The theatre state in nineteenth-century Bali*. Princeton University Press.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. Sage.
- Grek, S., & Lindgren, J. (Eds.). (2015). *Governing by inspection*. Routledge.
- Grek, S. (2014). OECD as a site of coproduction: European education governance and the new politics of policy mobilization. *Critical Policy Studies*. <https://doi.org/10.1080/19460171.2013.862503>
- Gustafsson, J.-E. et al. (2017). *Samling för Skolan*. www.regeringen.se/498092/contentassets/e94a1c61289142bfbcfd54a44377507/samling-for-skolan---nationell-strategi-for-kunskap-och-likvardighet-sou-201735.pdf. Accessed 4 May 2017.
- Haas, P. M. (1992). Introduction: Epistemic communities and international policy coordination. *International Organization*, 46(1), 1–36.
- 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>
- Heclo, H. (1974). *Modern social politics in Britain and Sweden: From relief to income maintenance*. Yale University Press.
- Inter-Agency and Expert Group on the Sustainable Development Goal Indicators (IAEG-SDGs). (2015a, June 1–2). *First Meeting of the Inter-Agency and Expert Group on the Sustainable Development Goal Indicators New York*. ESA/ST/AC.300/L3. United Nations.
- Ladrech, R. (1994). Europeanisation of domestic politics and institutions: The case of France. *Journal of Common Market Studies*, 32, 1. <https://doi.org/10.1111/j.1468-5965.1994.tb00485.x>

- Lahn, B., & Sundqvist, G. (2017). Science as a “fixed point”? Quantification and boundary objects in international climate politics. *Environmental Science & Policy*, 67, 8–15.
- Lewis, S., & Hogan, A. (2019). Reform first and ask questions later? The implications of (fast) schooling policy and ‘silver bullet’ solutions. *Critical Studies in Education*, 60(1), 1–18. <https://doi.org/10.1080/17508487.2016.1219961>
- Louis, M., & Maertens, L. (2021). *Why international organisations hate politics: Depoliticising the world*. Routledge.
- Lowndes, V. (2010). *The institutional approach in “theories and methods in political science”* (D. Marsh & G. Stoker, eds.) (p. 65). Palgrave.
- Martens, K. (2007). How to become an influential actor—The ‘comparative turn’ in OECD education policy’. In K. Martens, A. Rusconi, & K. Lutz (Eds.), *Transformations of the state and global governance* (pp. 40–56). Routledge.
- Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. University of Chicago Press.
- Meyer, C. O. (2005). Convergence towards a European strategic culture? A constructivist framework for explaining changing norms. *European Journal of International Relations*, 11(4), 523–549.
- Miller, P. (2001). Governing by numbers: Why calculative practices matter. *Social Research*, 68(2), 379–396.
- Newman, J. (2012). *Working the spaces of power: Activism, neoliberalism and gendered labour*. Bloomsbury.
- Nóvoa, A., & Yariv-Mashal, T. (2003). Comparative research in education: A mode of governance or a historical journey. *Comparative Education*, 39(4), 423–438.
- Nusche, D. et al. (2011). *OECD reviews of evaluation and assessment in education: Sweden 2011*. OECD Reviews of Evaluation and Assessment in Education, OECD.
- OECD. (2015). *Improving schools in Sweden: An OECD perspective*. OECD.
- OECD. (2016). *Reviews of national policies in education: Netherlands 2016*. <http://www.slideshare.net/OECD/EDU/reviews-of-national-policies-for-education-netherlands-2016>
- Olsen, J. P. (1998). Political science and organisation theory: Parallel agendas but mutual disregard. In R. Czada A. Heritier, & H. Keman (Eds.), *Institutions and political choice*. VU University Press.
- Porter, T., & Webb, M. (2004). *The role of the OECD in the orchestration of global knowledge networks*. Paper Prepared for the International Studies Association Annual Meeting, Montreal 2004.
- Prais, S. J. (2003). Cautions on OECD’s recent educational survey (PISA). *Oxford Review of Education*, 29(2), 139–163.

- Sabatier, P., & Jenkins-Smith, H. (1993). *Policy change and learning: An advocacy coalition approach*. Westview.
- Simola, H. (2005). The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative Education*, 41(4), 455–470.
- Stone, D. (2005). Global knowledge networks and international development: Bridges across boundaries. In D. Stone & S. Maxwell Global (Eds.), *Knowledge networks and international development: Bridges across boundaries* (pp. 89–105). Routledge.
- Stone, D. (2017). Understanding the transfer of policy failure: Bricolage, experimentalism and translation. *Policy & Politics*, 45(1), 55–70.
- Swedish Government. (2015). *OECD presents Swedish schools review*. <http://www.government.se/press-releases/2015/05/oecd-presents-swedish-school-review/>. Accessed 27 March 2017.
- UIS (UNESCO Institute for Statistics). (2016). *Laying the foundation to measure Sustainable Development Goal 4*. Sustainable development data digest. UNESCO Institute for Statistics. Retrieved 9 June 2020 from <https://unesdoc.unesco.org/ark:/48223/pf0000245559>
- UN. (2020). *Tier classification for global SDG indicators as of 17 April 2020*. United Nations Statistics Commission (UNSC). Retrieved 9 June 2020 from https://unstats.un.org/sdgs/files/Tier%20Classification%20of%20SDG%20Indicators_17%20April%202020_web.pdf.
- 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.
- Waldmüller, J., Jamali, H., & Nogales, N. (2019). Operationalizing Sustainable Development Goals in vulnerable coastal areas of Ecuador and Pakistan: Marginalizing human development? *Journal of Human Development and Capabilities*, 20(4), 468–485. <https://doi.org/10.1080/19452829.2019.1666810>
- Weick, K. (1995). *Sensemaking in organisation*, Sage.
- Wood, M., & Flinders, M. (2014). Rethinking depoliticisation: Beyond the governmental. *Policy & Politics*, 42(2), 151–170.
- Zito, A., & Schout, A. (2009). Learning Theory reconsidered: EU integration theories and learning. *Journal of European Public Policy*, 16, 1103–1123.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Beyond Objectivity? Storytelling and Reflexivity as Expert Work

1 INTRODUCTION: STORIES AND REFLEXIVITIES IN GLOBAL EDUCATION GOVERNANCE

In the introductory chapter of this book, I discussed how the pressures to make global public policy more inclusive, open and democratic have profoundly affected how experts work in these settings in order to achieve change. As the infrastructure of global education governance becomes more and more extensive and multi-polar, the efforts to quantify and commensurate vastly different geographical and political spaces stray further and further away from constructing ‘perfect’ statistical environments; what is more, apart from the technical difficulties that amount, the sustained effort to produce national and local statistical systems with the input from and for the benefit of the local communities make the work of International Organisations’ (IOs) experts more difficult, facing challenges both upstream and downstream, with many of them feeling that they have an impossible job at hand.

Moreover, although data-led governance has often been depicted as technocratic and ‘de-dramatized’, since the mid-2000s the ‘behaviour change agenda’ (Jones et al., 2013) emerged as a catch-all term for adopting a behavioural science approach that understands the production of public policy as combining cognitive and non-cognitive elements that include ingrained biases and affective ‘moods’. Scholars have thus paid attention to how these novel forms of affective expertise ‘re-humanises’

the policymaking process, by establishing the need for policymakers to design policies which are emotionally informed. There is burgeoning literature on the affect theory and the ‘emotional turn’ (Hoggett & Thompson, 2012). Similarly, feminist geographers have written persuasively about the ‘new enthusiasm for an emotionally attuned approach to government’ (Pykett et al., 2016, p. 1). Therefore, we see increasing tendencies to envelop number with stories that can help to better make sense and produce affective responses to the issues at hand.

Another more recent direction of contemporary public policy literature is the increased focus on democratic innovations. (Bua, 2019, p. 282). Though still in its infancy, this scholarship has suggested that the horizontal, non-hierarchical character of the ‘governance turn’ (Bevir, 2011) in fact coincided with profoundly exclusionary and undemocratic practices, such as the colonisation of public policy by private interests (Lee et al., 2015), the depoliticization of social issues (Streeck, 2013) and the centralising and managerial tendencies of contemporary public administration (Gamble, 1994). The proponents of democratic innovations suggest that inclusivity and a consideration of a variety of evidence-bases, including non-quantitative ones, can enhance and expand the epistemic basis of decision-making. Similar with other fields, education governance has been guided by developments in the field of participatory governance and democratic innovations in an effort to explore the potential of data storytelling to create common, democratic spaces of deliberation through the use of visualisations.

Thus, the key question that arises in this chapter is how education experts try to remain faithful to their epistemic credentials as purveyors of factual knowledge, while also using instruments like stories and visuals in order to persuade actors to participate in these contentious processes. In trying to answer this question, this chapter turns to an examination of storytelling and reflexivity as key tools in the production of expert knowledge in this new era of ‘democratising’ global education monitoring. In this context, reflexivity is not only a thought process that is necessary for the work of expert actors in order to continuously make sense of what their work involves, to justify its failings, or to explain their moral predicament. Such a view of being reflexive to the applied and contextual character of producing knowledge for policy became a key feature in knowledge production work from the 1990s onwards, as it has been manifested in the Mode 2 literature. More recently, however, we see the workings of these ‘softer’ persuasive instruments, such as stories

and reflexivities not only as an epistemological endeavour but also as a key political *resource*: as this chapter will show, expert actors use reflexivity as a political tool in their efforts to construct consensus and mobilise the participation of countries and their representatives in monitoring agendas and frameworks.

Therefore, although previous literature has seen reflexivity as an essential tool for scholars to address the ways that reflexivity may affect the research we do, here I am interested in the ways that reflexivity is a necessary tool for analysing the work of experts, as well as the socio-material manifestations of expert work, such as the case of data visualisations. The following section will draw upon the theorisation of the relationship of quantification with *qualification*, a key process as I will show, in understanding how the work of numbers is enveloped and made sense of through the application of value judgement. The chapter will then move to the discussion of two empirical examples from the field of the global governance of education: first, the use of interactive data visualisations as storytelling devices, and second, the role of instrumental reflexivity—i.e. the ways actors may practise reflexivity as a way of creating consensus and achieving commensurability.

2 EXPERT REFLEXIVITY: STS AND THE WORK OF QUALIFICATION

What do I mean by a relationship of stories and visuals with the production of evidence-based policy? Examples of such interactions abound: in fact, the more collective, critical and fundamental the policy issues, the bigger the influence of stories and visuals in contributing to the shaping of policy problems (Mesch, 2013). From the historical cases of the American civil rights movement, to feminist art, all the way to the global challenges of sustainability, migration and public health, stories and visual imagery have always been actively present in shaping the formation of new public arenas. Although the contemporary hegemony of data-led decision-making is rooted in rationality as the Enlightenment's promise of knowing and governing the human condition, the rise of the modern scientific age never rejected stories and images as a way of observing, recording and transforming the world. On the contrary, science and art have always been closely entangled. Science was not merely about producing universal and objective knowledge; it was also spectacular, in

the literal sense of the term: theatres of science were venues for instruction, but also entertainment and social recognition (Blatchford & Blyth, 2019). Although political science and policy studies have developed an interest in narrative and visual approaches for the understanding of the relationship of knowledge with policymaking fairly recently, the field of Science and Technology Studies (STS) has always seen visual practice as a key medium via which new forms of knowledge, methodologies and engagements can be generated.

STS originated in the 1960s when critical debates on the societal role and impact of scientific and technological innovations emerged. From Edinburgh's 'Strong Programme' of the '70s and '80s (Bloor, 2013), to Latour and Woolgar's ethnographic study of the implicit, tacit knowledge and embodied skills that scientists develop (1979), and to Mol's 'ontological multiplicity' (2002), the centrality of story/image-making in STS is not coincidental: it is routed in STS's research agenda that focuses on technology and materiality; boundary work; subjectivity and the senses; and embodied, situated and enacted forms of cognition (Benschop, 2009). Indeed, John Law suggested that 'knowing and its methods are materially complex and performative webs of practice' (Law, 2017, 47). Law gives examples of art/science collaborations as 'hybrid knowing spaces' that work through performance, text and simulations in achieving knowledge 'that might be otherwise' (Law, 2017, p. 48).

Further, the STS concept of 'qualculation' (Callon & Law, 2005) demonstrates the very fine balance between calculation and judgement. According to Callon and Law (2005), complex decision-making requires both acts to be performed together. There are no instances where a mere calculation can give the answer to a wicked issue, given that even the very act of calculation itself requires the application of judgement. According to Moser and Law, qualculation 'is a way of drawing attention to the fact that the two (which are habitually treated as being different in kind) both become possible—indeed they are only possible—because they array and manipulate appropriate elements within a single relevant frame in order to achieve an outcome or a conclusion' (2006, p. 66).

Perhaps the body of scholarship that theoretically guides this analysis most powerfully is the emergent field of Art and Science and Technology Studies (ASTS) (Borgdorff et al., 2020). STS has worked closely with artistic research, a field of studies that explores the production of knowledge and research through or in artworks. According to Borgdorff et al., 'artist-scholars in this field focus on the knowledge, understanding, and

experiences enacted in creative processes and embodied in artistic products such as artworks, compositions, and performances' (2020, p. 1). Here I draw on Caroline Jones' and Peter Galison's (2014) *Picturing Science, Producing Art*, which richly demonstrates how art and design are deeply entangled with socio-technical worlds. The authors suggest that 'what much of the focus on "art" and "science" as discrete products ignores are the commonalities in the practices that produce them. Both are regimes of knowledge, embedded in, but also constitutive of, the broader cultures they inhabit' (Jones & Galison, 2014, p. 2).

Such observations need not only be made by artists and scientists however: similarly, METRO fieldwork showed how education policy experts persistently describe their daily business as resembling artistic practice more than scientific work: they suggest that their day-to-day job 'is more of an art, than a science' (METRO interviewee). Although METRO focused on the role and effects of quantification in global governance, this—albeit cliché—phrase led to a reconceptualization of education expertise, in order to include other instruments and tools that facilitate the production of knowledge for governance. Borrowing on STS, what are the boundary knowledge spaces onto which policy-makers build upon in order to make decisions? When/how does a data visualisation, for example, become an object of beauty *and* information? When does the photograph of a drowned child in a Mediterranean beach become an artwork *and* a leverage point for policy change? When does a film, like the recent *Don't Look Up* (2021, by Adam McKay), become popular both as black comedy *and* as a stark warning to our common predicament? Exploring the production of stories and visuals as a space of knowledge and policymaking capitalises on the decades long study of the entanglement of art and science by STS in order to examine if and how policymakers turn to art to develop techniques, values and skills that may be less tangible than statistical data, but still a crucial, yet so far unknown, part of the policy repertoire.

Therefore, in order to make sense of the production of education expert knowledge, I will now turn to the theoretical underpinnings of the concept of 'qualification' (Reinicke, 2015), the process via which actors make value judgements on the basis of the decisions and choices they are confronted with. These value judgements might not necessarily take into account pre-conceived categorisations, classifications or even other expert advice. In fact, such value judgements are seen as being made continuously, given the infinite world of commodities and services available:

selecting a lawyer is, for example, a decision perhaps not only based on the *value* of the services that may be on offer, or on the ranking of the local solicitors' performance, but also on other *values*, too, such as trust, personal acquaintance, fame or respect.

In other words, decisions on many aspects of everyday life are not only dependent on statistical knowledge (that tends to standardise in order to reduce multiple values in a specific value: the process of *quantification*). Rather, they are based on judgement of the decision's (or the good's) values (the process of *qualification*): this is a process that, instead of standardisation, requires a process of 'individualisation' (Callon, 2002, p. 267). Translated into the context of global education governance, and despite the prevalent focus on analyses of quantitative expertise as a process of commensuration and standardisation (including the author's), experts in the field are continuously confronted with the very specific ('individualised') challenges and values of local populations. Although making judgements is an inherent aspect of the production of quantification, the process of qualification denotes more than that: it is the process whereby certain *measurable and standardised values* (in the statistical sense) are being consciously opened up to assigning *certain political values* to the good in question (or they establish new 'orders of worth', following Boltanski & Thévenot, 2006).

In the increasingly dispersed governing space of the global education policy field, such a distinction between quantification and qualification, albeit thin and transient, is crucial to understanding the ways experts negotiate their epistemic capital with the political values on the ground, as well as their own personal ones as they go about their day-to-day work. To clarify—and return to the analysis earlier in this section—my analysis is not confined to the tensions of Cochoy's 'qualculation' (2008) that all quantification practices involve judgement. Calculation does not grow on trees, as Callon and Law suggest (2005): it requires time, money and effort and the sociology of quantification has given us persuasive accounts of the judgements inherent in all quantitative practices (Strathern, 1987). Rather, the focus in this chapter is on storytelling and visualisations as the socio-material devices that often assist experts and policy actors to be reflexive and persuasive, so as to engender trust, optimism and confidence that datafication of education governance is the only way to achieve educational success and equity.

This, in some ways, is the reverse process of quantification, via which *values are ascribed to value*: within the field of global education governance, qualification has become a key component of consensus-building and of increasing participatory and inclusive decision-making practices. Post-PISA and the global acceptance of the orthodoxy of datafication in education policymaking, this is the complex task education experts are asked to deliver, as they try to always match global processes of commensuration with local struggles over priorities and political ideas.

The concept of qualification, as I will show, helps establish and analyse the role of personal and collective values in the struggle over establishing conventions of worth: facilitated via stories and images, reflexivity is a key resource, both at identifying and codifying values at the level of the individual expert/actor (values that make their work, however utopian, worth doing), as well as at the level of working through local political values and agendas and trying to ‘marry’ them with the more top-down global goal-setting. Here, the METRO project found that in the context of the production of global metrics, reflexivity is not only a thought process that is necessary for the work of experts to continuously make sense of what their work involves, to justify its failings, or to explain their moral predicament. Rather, reflexivity is also a key political resource: experts use reflexivity instrumentally and as a political tool in their efforts to construct consensus and mobilise the participation of countries and their representatives in monitoring agendas and frameworks. The remainder of the chapter will explain the socio-materiality of visualisations and stories in enabling reflexivity, as well as the instrumental use of reflexivity by experts in order to emerge as the trusted partners of the Global South.

3 PLAYING GOD: EDUCATION DATA VISUALISATIONS AND THE ART OF WORLD-MAKING

We know by now that the visualisation of measurement facilitates the understanding of complex information sets and supports interpretation and sense-making; more than that, there is increasing realisation that visualisations prompt engagement with calculative technologies (Gatzweiler & Ronzani, 2019; Quattrone, 2017). In particular, the properties of data visualisations endow them with an aesthetic appeal that affects how users interpret them, appropriate, and make meaning with them (see Espeland & Stevens, 2008; Kornberger, 2017). For these

reasons, scholars have called for a reconceptualization of data visualisations in the digital age, contending that the visualisation of ranked performance can act as a persuasive and robust ‘judgement device’ (Begkos & Antonopoulou, 2020). For example, recent research shows how higher education is influenced by the rankings’ visual formats, which ‘allow platforms to at once display cascades of inscriptions in a pleasant, aesthetic manner and further complicate the numerical-ordinal basis of traditional ranking systems’ (Decuyper & Landri, 2020, p. 12).

The interpretative flexibility of data visualisations makes them prime sites to explore some of their generative effects (Pollock & D’Adderio, 2012), including generating the possibilities for reflexivity, as I will show below. Visual elements are not only important because they support calculations but also because they offer interpretative clues that cognitively and aesthetically engage with the users of data (Espeland & Stevens, 2008; Quattrone, 2017). Thus, data visualisations are critical to knowledge brokerage, as they enable the communication of research findings to different discourse communities and play important roles in the legitimization and dissemination of data production (Allen, 2018). As I will explain, increasingly, data visuals not only substantially assist with the communication of data: they further enhance the data’s influence, as they facilitate a diversity of interpretations, translations and ultimately the reflexivity of those working with them.

3.1 *‘No-One Left Behind’—Data Storytelling as Reflexivity-Making?*

Each of us walks around with a bunch of stories in our heads about the way the world works. And whatever we confront, whatever facts are presented to us, whatever data we run into, we filter through these stories. And if the data agrees with our stories, we’ll let it in and if it doesn’t, we’ll reject it. So, if you are trying to give people new information that they don’t have, they’ve got to have a story in their head that will let the data in.

This section will examine storytelling as an increasingly popular form of visualisation in the education and development world. Storytelling in public policymaking has emerged in recent years as a powerful tool for policymakers and researchers to communicate complex messages in order to reach larger audiences. Either used as a knowledge brokering tool in

negotiations among policy actors or weaponised as an advocacy medium in activism, visual storytelling uses the essential elements of story-making across time and space: it is comprised by main characters, a setting, a plot and a moral, in order to help make causal relationships apparent and to frame ‘facts’ and data within particular narratives.

Crucially, the aim of visual storytelling, as I will see below, is less about communicating specific data fast. Rather, it relates to the making of larger frames of political values, where data, numbers and performance monitoring via country rankings, are only one of the building blocks of data ‘world-making’. Instead of rational and objective, visual storytelling is wholly interpretivist in nature and function. Despite the appearance of an objective rationality purported by numbers, stories are meant to be used as tools of reflexivity and data translation. Their function is to construct the narrative frame within which a carefully selected data pool can offer objective comparative country and regional performance. At the same time, however, the comparison is carefully massaged and shaped in a way that a main problem is addressed, key challenges discussed and—usually—some solutions offered.

Data storytelling is particularly interesting for the analysis of knowledge production for governing. Instead of concealing the inbuilt biases and assumptions that all objectivity-making requires, it does precisely the opposite. That is, it works with people’s engrained world views and attempts to shape and reshape them by pressing towards the making of new political problems and political values. 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. This is not ‘facts versus values’ evidence-making; the effect is, in fact, almost antithetical to the cold rationality of statistical numbers. Data storytelling uses facts *for* value-making, and in doing so exploits the subjective and contingent nature of knowledge-making.

*Left Behind*¹ focuses 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

¹ <http://uis.unesco.org/apps/visualisations/no-girl-left-behind/>.

data from UIS. The visual focuses on the gender inequality problem, and in particular the non-participation of African girls in education (Fig. 1).

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. The image therefore follows a very common stylistic feature found in art; that is, it creates a sense of exoticism. In doing so, through the subtle connotations which align this one with numerous other stories about worlds distant from the West, the image has already served towards framing this story within well-known and classic art historical framings of picturing the ‘Other’. These are not just any schools, any girls or any countries: this is Africa.

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



Fig. 1 Front webpage of Left Behind visualisation

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 visual, 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, 0.3 Persistence of Illiteracy among Women, 0.4 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. Although 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 performance is completely abandoned here. Although there are better and worse country cases (this is the function

of any graph and therefore of these graphs, too), 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 (Fig. 2).

An analysis of those data visuals immediately reveals a range of differences and similarities: there is a balance of change and stability. Clarity is paramount. There is no alarmism, although negative performance is being reported, too. Although the main character remains the same (i.e. African girls, women or teachers), the plot is very carefully crafted in order to move from setting the context (0.1 The Last Mile: ‘there are good news...but the gender gap persists’), to a discussion of all challenges (in ‘Chapters’ 2, 3, 4) to the relatively uplifting final section on the necessity to have a 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’ (Fig. 3).

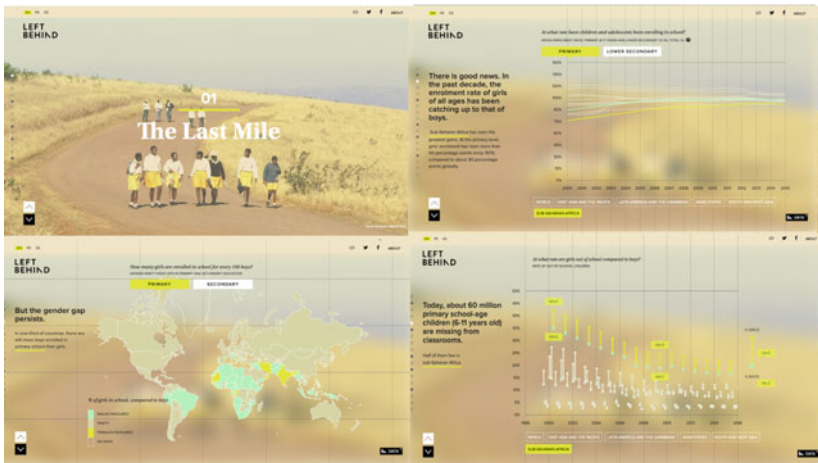


Fig. 2 Snapshots of Left Behind visualisation (01)



Fig. 3 Snapshots of Left Behind visualisation (02)

This is perhaps the first step in constructing actionable knowledge: enlist one's audience not only to read and understand, but also to share their experience of the African girls' education story and mobilise others. Interestingly, the visual does not do 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. 4): 'What do you think it will take to leave no girl behind?'

This question in many senses is at the crux of this chapter's argument: rather than finish off with a definitive memorable statement, or a killer graph, apt for describing the severity of the issue, *Left Behind* ends with inviting 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 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

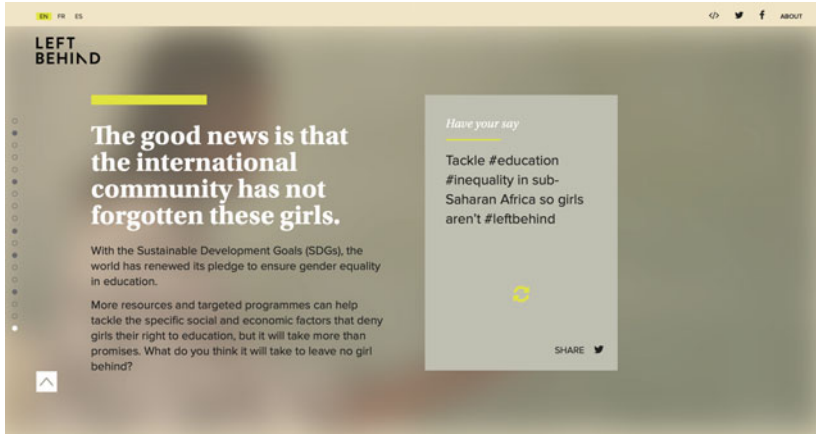


Fig. 4 Left Behind last page

worlds that data visualisations fabricate, worlds into which specific and precise policy facts do not matter as much as the reflexive possibilities data (and especially an effective visual data story) can open up.

Left Behind is an illustrative case of the power of numbers, combined with images and storytelling, to communicate and persuade. The next section describes other tools for enhancing and further spreading the legitimization of quantitative expertise: this is reflexivity and its instrumental use by actors who are in charge of processes of engendering trust and collaboration between local politics and actors with international monitoring agendas.

4 INSTRUMENTAL REFLEXIVITY AND EXPERT WORK

As the previous section discussed, data visuals facilitate reflexive practice, allowing for multiple translations of the numbers presented in them, as they work alongside images, colours and stylistic features that work with users' interpretive repertoires. In this section, I will move this analysis one step further, in order to explain how reflexivity has become a key resource, not only in the interpretative work that numbers require, but also as a political instrument: it is being foregrounded and used as the main means of constructing and maintaining relationships of trust between experts

and countries. In this way, experts can instrumentalise reflexivity for political action (for a broader discussion of the different uses of reflexivity in global public policymaking, see Bandola-Gill et al., 2023).

In order to illustrate the ways in which experts mobilise instrumental reflexivity, I focus on one empirical example exploring the expert work of the UNESCO Institute of Statistics (UIS), not despite, but *because of* their explicit and intentional reflexive accounting of the challenges of producing quantification for the benefit of countries in the global South.

In order to contextualise the case, we need to understand that the history of the construction of the SDG4 is one of struggle. As already discussed in previous chapters, the two main opposing camps were, on the one hand, the 'Education For All' (EFA) movement, and on the other, the process of work undertaken as part of the Millennium Development Goal education indicators. For reasons of brevity and in order to avoid repetition, I won't develop this history here, but simply state that the two groupings had very conflictual views about the best measurement approach in education to be undertaken: EFA pushed for a diverse set of goals that would acknowledge a broader, humanistic approach to education, whereas the MDGs education experts wanted to find a much more specific and measurable set of instruments, favouring a utilitarian view of education and focusing on key metrics such as literacy and numeracy.

Therefore, in the face of the threat of an education-specific goal being excluded from the SDGs due to the inability of the two groups to find common ground, a solution was found and the worst was avoided: the compromise led to the production of the SDG4. Nonetheless, even if the contestation seemed to temporarily abate, it never really went away. On the contrary, the continued challenges of meeting the SDG4 goals and constructing a solid set of indicators to do so have intensified the struggle and conflict in the field. It is in this space of clash that UIS managed to emerge as the reflexive and trusted international organisation, distinct from others who are seen as more technocratic and representing interests of the Global North. Having worked in countries of the Global South for decades, the UNESCO Institute of Statistics (UIS) was the expert organisation with long-standing links and relationships with the relevant countries, as well as the ability to use data failings (and often of their own making- UIS had had some serious measurement project failings in the past, for detailed analysis see Fontdevila 2021) in order to advocate for the notion of accepting the production of 'good-enough', (rather than

precise) data, and the political (rather than purely technocratic) uses of target-setting for coalition-building and agenda-setting.

However, how did UIS manage to maintain their position as a data producer alongside powerful others, while at the same time appear to be at the side of struggling countries? First of all, UIS adopted a much more practical rather than ‘perfectionist’ approach to the production of global metrics in education. Instead of advocating for a single measurement tool (like the other IOs did), they focused their efforts towards accommodating the use of different assessments and harmonisation methods. In contrast to other actors, such as the World Bank or the OECD that would have been much stricter in the choice of method (with a preference for their own instruments), the UIS developed more of a ‘patchwork’ approach: they recombined several already available and legitimate models, recognising openly the limitations of each and emphasising the potential for complementarity. Following this more pluralist method, they appeared a lot more accommodating in their data demands, while acknowledging the challenging circumstances that many countries face, trying to collect both commensurable data for the top-down demands, and dealing with the day-to-day requirements of producing governing data. Due to this particular stance that UIS adopted, many interviewees recognised it as perhaps not a data superpower, but as the trustworthy actor that recognised the unequal character of the data production market and thus the difficulties of creating an inclusive space, with the emphasis on the principle of country ownership.

Second, perhaps more importantly, UIS, primarily through its outspoken Director, Silvia Montoya, publicly discussed the imperfect character of global learning data, as well as the political nature of the indicator process. In doing so, she used a highly reflexive approach, emphasising the epistemic challenges of comparison across highly different contexts, as well as the need to find better approaches to coping with missing data and the deep inequalities in statistical capacity of the participating nations. Under Montoya’s openly reflexive leadership, UIS nurtured types of approaches for the collection of data that are hybrid, and brought together different types of assessments, insisting that the different data alternatives are not mutually exclusive but reinforce one another. More importantly, this incremental approach went against selecting one specific method as technically superior to others, and thus was politically much more in tune with countries and their specificities.

Thus, not only a middle-way forward was found, but countries also felt respected for their context-specificities and were not sidelined:

There has been significant growth and improvement in the field of learning assessment across the world. Yet today, it is impossible to provide a global perspective of what children are learning... **We must be pragmatic.** As explained in previous blogs, the best measures and methodologies in the world will amount to little if countries cannot produce them. We must therefore take a pragmatic approach, which may mean mixing the options. This stepping-stone approach was widely endorsed by stakeholders attending the June meeting. They understand the political stakes, the technical issues and the need to find a balance between pragmatism and accuracy...

We need to recognize that SDG 4 indicators are barometers – showing which countries (and, for equity’s sake, ideally which segments of which countries) are making progress and which countries need help. Instead of aiming for the most technically rigorous methodologies, we may better serve the world by taking a pragmatic approach to producing the global measures while helping countries improve the quality and use of their national data. (Montoya, 2017; my emphasis)

As is apparent in the above quotation, UIS used reflexivity *instrumentally* to reaffirm and strengthen its authority in the education measurement realm as the only trustworthy, ethical and transparent expert broker. Montoya’s reflexive account is not limited to an assessment of the epistemic limitations of the monitoring exercise. Instead of approaching the construction of indicators as a purely technical exercise, despite its apparent limitations, the UIS openly discussed the *political* nature of the debate as well as the vested interests that shaped it (for example, its director exposed the inefficiencies of the ‘learning assessment market’ in two influential blogs in 2019²). UIS openly admitted that there is no perfect way of doing this kind of work and that technical rigour would have to go hand in hand with a more pragmatic approach: this way, reflexivity became the prime instrument for the organisation to bolster its credibility and create minimum consensus in the field. As a consequence, the notion of ‘good-enough’ data gained centrality, as the political choices

² <https://gemreportunesco.wordpress.com/2019/04/26/the-learning-assessment-market-pointers-for-countries-part-1/>.

and judgements were not hidden but in fact, displayed publicly and used repeatedly in talks and all sorts of public fora.

Thus, similarly to the use of data storytelling, the concept of instrumental reflexivity describes the considerations experts engaged themselves in in the cases where the epistemic qualities of quantification (objectivity, de-contextualisation, universality) were in tension with the political goals of measurement. Here, not only experts did not avoid exposing the political nature of numbers, but also even went as far as to mobilise and instrumentalise it, in order to achieve their goals. Of course, one has to take into account the interdependencies, competitions and collaborations between IOs in order to get a fuller picture of how IOs interact and assume different, complementary identities as they work collaboratively: while some may take the high ground and defend their authority by sticking closely to its objectivity and trustworthiness, others choose to benefit from getting their hands dirty and muddle through political contestations and imperfect numbers.

5 EXPERTS' REFLEXIVITY IN GLOBAL EDUCATION GOVERNANCE: THE ROLE OF VISUALS AND STORIES

Although policymaking has always been imbued with visual messaging, the visualisation of political communication became particularly central with the rise of data-driven governance. As we have experienced during the last decade, the acceleration of datafication of contemporary policymaking has closely been accompanied by the rise of data visualisations as a key mode of not only political communication, but also policymaking itself (cf. Bekkers & Moody, 2014; Amit-Danhi, 2021).

As the empirical analysis of the *Left Behind* data visualisation showed, data visualisations are effective communication tools, as they have the ability to minimise the complexity of represented issues and summarise them for multiple audiences. As argued by Falisse and McAteer (2021), the success of specific data visualisations relies on their ability to simultaneously summarise complex information and contextualise them within the broader policy context. This quality positions them as 'boundary objects' (Star & Griesemer, 1989) mediating and brokering between different communities; in the case of *Left Behind*, they work at the inter-spaces of policy production, accommodating the interests and needs of both IO experts and local policy communities. Thus, visualisations have

become vital passage points in complex socio-technical systems, as they are located in-between different forms of networks (Rose et al., 2014).

As I have shown in previous work, data visualisations do not only target but also can outline and constitute groups of stakeholders around issues, as they act as ‘alignment devices’ that orient diverse actors towards a common goal (Bandola-Gill et al., 2021). In the case of *Left Behind*, national performance appears to be of less importance than the need to create alignment and consensus around the need to reduce gender disparities in education; the visual promotes reflexivity around these issues, represented in such a way as to avoid the ‘winners and losers’ older narrative of comparative performance and instead enhance the ideas around universality and common purpose. Indeed, although the underpinning rhetoric of data visualisations is one of political neutrality and technocracy as they are deemed to be representing ‘facts’ (Kennedy et al., 2016), their effects are in reality more political: data visualisations are effective tools of persuasion (Pandey et al., 2014) and they carry this function through multiple means. On the one hand, data visualisations are the most explicit sites of the politics of visibility and invisibility of numbers and data (Espeland & Lom, 2015; Espeland & Yung, 2019)—they make some aspects of the visualised problems evident, while others disappear. More importantly, and as we have seen in the *Left Behind* visual, they offer not only description but also the possibility of reflection and interpretation (Bekkers & Moody, 2014a; 2014), as well as emotion (Kennedy & Hill, 2018; Lefsrud et al., 2020). Data visuals are *performative*—they do not just reflect the represented phenomena but construct them and their fields of practice. Even though this process is often considered implicit and almost automatic, at times it can be consciously mobilised by actors aiming to advance their political agendas (Fileborn & Trott, 2021).

Similarly, experts working in IOs shared faith in numbers to bring transformative change, but were also acutely aware that their work is mostly political (Bandola-Gill, 2021). More importantly, they were happy to reflect on, discuss and share the challenges of their day-to-day work. Providing expert advice was seen as a process that required a specific set of qualities that did not limit themselves to quantitative expertise. On the contrary, expert qualities needed to be a lot more diverse, empathetic, creative and adaptable: they involved an understanding of data but also of the local contexts; humility and perseverance in the face of limited funding and the diversity of interests and value-systems; an ability to foresee change and place themselves at the best possible place to tame it;

and finally, the skill to transform a perceived obstacle (the lack of precise data) into a valuable instrument for advocacy and consensus-making (the concept of ‘good enough’ data) (Grek, 2020).

In this context, the co-existence of these different epistemic orders (related to the quality of data and the politics of measurement), political orders (related to bringing actors ‘on board’ and producing contextualised measures) and value orders (related to different ethical priorities and cross-cultural ways of working) meant that the work of the experts went beyond just ‘producing numbers’. This multifaceted navigation between different priorities required them to mobilise different styles of knowing. Reflexivity, therefore, emerges as one of the new skills in the expert arsenal. As such, it is both an epistemic practice (as traditionally discussed in the literature on the topic) but also a practical and strategic tool that can be mobilised in the context of complexity.

Approaching reflexivity as *practice* allows for unpacking its core elements. First, it helps analyse not only the practices of experts themselves but also the socio-material tools that support and promote reflexivity for wider audiences and users—in this view, reflexivity in education governance becomes a *shared* resource rather than individual endeavour only. In this regard, and as I have shown in this chapter, the *visualisation* of data is key, as data visualisations appear to perform a significant function: they work towards the political goal of aligning policy priorities towards specific global challenges, many of which might look similar, yet, they can also be adjusted because contextual and regional specificities and trajectories render them different. This multiplicity and ‘adjustability’ does not take away from the authoritative nature of the data. On the contrary, it further reinforces data credibility by making them relevant to all without antagonising lower-performing countries. Therefore, we observe how socio-technical devices materialise instrumental reflexivity, as not only as a practice applied by experts themselves, but also as a way of creating the conditions for promoting data reflexivity for users, too: a certain kind of ‘world-making’ that only metrological realism could have enabled and promoted. In so doing, data visualisations allow their users to create acceptable narratives around both their own specific country performance and the common global sustainable development agenda. Data visualisations retain the illusion of the political neutrality of their producers foregrounding an ‘issue-based’ message (for example, gender equality) rather than focusing on augmenting competition and

peer pressure to achieve individual goals (e.g. benchmarking country-level performance). As such, data visualisations create reflexive spaces that align actors with diverse interests and interpretations of performance by allowing for the coexistence of multiple, often contradictory interpretations of data.

Further, as we have discussed in detail in previous chapters of this book, IO experts work in an institutional and political context that is shaped by ever-expanding infrastructures of measurement (Merry, 2019) and pressures to decolonize global governance (Rottenburg, 2009). This context is not only highly fragmented by the growing number of actors, but it is also increasingly participatory and requires constant mediation and brokerage (Grek, 2020). Furthermore, an increasingly pressing issue is the development of measurement approaches that avoid alienating low-performing countries by continuing to pressurise them to conform with ‘best practices’ from elsewhere. International organisations and other key global players purport to design their measurement programmes following equity paradigms, where all the countries—and especially the developing ones—are seen as leading on tackling the global challenges (Best, 2014).

Against this backdrop, the work of reflexivity ‘softens’ the rigid measurement of data and facilitates the promotion of participatory paradigms of global governance. Data, apart from objective, has to be reflexive, and therefore produced in a way that navigates political pressures while communicating the urgency of the global problems as truly global—affecting the entire international community. Using either data storytelling or applying reflexivity instrumentally and strategically, expert actors can entice participation in measurement programmes (Desrosières, 2015; Le Grand, 2003, 2007), while keeping competitive behaviours—that would be dysfunctional in the global policy space—at bay (e.g. a sense of zero-sum competition, gaming, cherry-picking or the manipulation of data—see Espeland & Sauder, 2007; Merry, 2016; Slager & Gond, 2020). Arguably, the global ‘need’ for quantification and performance measurement has never been as perceivably legitimate as it has been since the introduction of the SDGs framework. Simultaneously, there has never been as much attention paid to how global performance measurement may be a form of ‘southering’ (Grotlüschen & Buddeberg, 2020) that presents developing countries as regions of persistent deficit, under the surveillance of Western institutions through different forms of quantification (Arora, 2016).

It is within this context that ‘old’ formats of presenting data (graphs and tables) do not seem to be suitable to capture this multiplicity of values and needs. The qualities of interactivity, engagement and trust are essential in heterarchical and polycentric settings such as the global governance space, where different ‘hierarchies and orderings intertwine and reproduce, none of which can claim to be dominant or even to be fixed’ (Esposito & Stark, 2019, p. 15). Since no single order shared by all exists in such a space, expert work seeks to achieve more ‘equitable’ and politically acceptable solutions by ‘softening’ the data’s appeal through acts of reflexivity; this is achieved via multiple means and it might entail visuals and stories, as well as the inclusion of diverse methodologies in an effort to push for epistemic justice (ostensibly, at least) rather than division. Instead of ordering data according to performance, reflexive practices increase the visibility of areas of concern and potential intervention without seeking to ‘shame’ explicitly any country.

Technological developments during the last decade have greatly enhanced the possibility to offer more intricate analyses of global data, as well as allow the use of data in diverse and more creative—and strategic—ways. For example, through the application of interactive tools, data visualisations are not simply communicated to the user; rather, although the messaging remains focused and clear, the user is also given the tools to engage with the digital interface. By ‘playing’ with the interactive formats, the users can see the multiplicity of data, and choose to work with aspects of it that interests them more. What emerges at the end of this ostensibly playful interaction is a message that does not seek to identify out a clear ‘loser’ because—rhetorically—there is none. By offering multiple views of the data, as well as a range of other visual and stylistic tools, ranked countries become de-individualised and move towards being aligned. The interactive formats METRO explored are a clear departure from the traditional visibility that is perpetuated by more conventional data visualisation whose argumentative power and appeal are tightly linked to their capacity to communicate ‘winners and losers’ almost at a glance (Bevan & Fasolo, 2013; Pollock & D’Adderio, 2012; Wedlin, 2006). On the contrary, data storytelling in the context of global education governance capitalises on more subtle qualities: similar to what is often expected from experts themselves, visuals invite engagement, they afford personalization, and seek to adapt to individual preferences and priorities.

Indeed, this is the central role of reflexive numbers in this context: data has to be clear enough to point to problems and inspire collective action,

however, without shaming lower-performing countries. The instrumental use of reflexivity in this context actively *facilitates* this multiplicity of interpretations and fuels the messaging and its political acceptability. The inherent multiplicity does not take away from the authoritative nature of the visual and the data it carries. On the contrary, it further reinforces the credibility of data by making them relevant to all without antagonising lower-performing countries. In global governance, expert work uses reflexivity to re-adjust country monitoring, while offering a steer towards a very specific and discursively ‘universal’ set of goals.

This point on the inherent multiplicity of data visualisations leads to explore them as ‘world-making’ tools. As argued by Latour (1986), visualisations stabilise specific versions of reality; they can make impossible things realistic and make possible objects more probable than others. In the METRO findings, we explored how data visualisations not only allow for the exploration of multiple aspects of the data but also enable customisation that allows the user to choose different value dimensions in accordance to their own preferences. They are conceived and designed in a way that allows expert—but also their users in the broader sense—to be reflexive and create their own knowledge and interpretations.

Further, having interviewed over 80 experts working in international organisations in the fields of education, poverty and statistical capacity development, METRO analysed the reflexive accounts of these actors’ day-to-day business, as they went about describing and justifying their work. Reflecting on my own expectations of what these accounts might entail, we anticipated that interviewing them would require more intensive probing to get them to explain the limits and challenges of quantification and the types of political work required to successfully implement metrics. The project team did not underestimate them and never thought that their technical expertise would not allow them space to be analytical; our surprise did not relate to the fact that they were thoughtful and eager to reflect on what their work involves. What did surprise the research team was the extent to which, time and again, many of these actors treated the interview space as a cathartic zone, where they would freely share their exasperations at being asked to achieve the unachievable, but also a space where they would share their conviction that measuring inequality was the only available means not only to know, but also crucially to raise awareness of the injustices communities—in the global South in particular—have to endure and overcome. Reflexivity therefore was not only a

thought process experts exercised as part of the encounter of the interview; instead, they described it as a tool in their day-to-day job, as they were tasked to insert meaning to their work of numbers, to persuade and to build relationships of trust and reciprocity. In other words, apart from a focus on experts as the ones holding the epistemic capital to *know* global education governance by naming and measuring it, statisticians, IO experts, and national and local decision-makers reflect on their practices that produce the monitoring system and wonder—other times despair—over how much or how little real-world effects their work has.

Thus, reflexivity appears to be doing a lot more of the heavy lifting of quantification than the literature has so far discussed. As the case of UIS has shown, reflexivity is not merely a process of self-appraisal by experts, as they make sense of their work in an internal dialogue between their personal values and aspirations and their activities on the ground. More than *self*-reflexivity, I showed the ways actors used the process of opening up the black box of number-making not only to us as researchers, but also with those in the field—including colleagues, collaborators and even policymakers. Experts purposefully put reflexivity to work, in order to, on the one hand, explain and justify choices as they muddle through trying to establish some order in the messy realities of quantifying complex problems, and on the other, as they actively attempt to imbue data production with the political values of inclusivity. Thus, they purposefully apply processes of *qualification*, as almost the reverse process of quantification: in their efforts to engage and co-opt communities, they need to—momentarily, at least—move away from the rationality and objectivity of commensurability, in order to open up these numbers to contextualisation and even contestation.

Although seemingly antithetical to the production of quantification as the process through which multiple values come together and are expressed through their representation by a single value (the one that can then represent multiple realities and thus be commensurable), *qualification* is a sine qua non to quantification. This is not simply because judgement is inherent in every single decision, no matter how large or little, over the making of numbers (i.e. what the concept of qualification denotes). Although such considerations are important, one of the most startling METRO findings was that expert actors were happy to go as far questioning their own authority by opening up a debate about numbers, where theirs and their interlocutors' political and personal values would acquire almost the same weight as the data itself: in casting light on the

ways that reflexivity becomes an essential element of the performativity of qualculation, Skeggs eloquently suggests that ‘values will always haunt value’ (2014, p. 1).

I am not claiming here that experts at any point during their accounts of their work in the field questioned the use, validity and significance of their numbers’ work; quite the opposite, as their quotations eloquently show. What instrumental reflexivity shows is that in order to make quantification work, experts need to re-attach political values to numbers, and thus allow them to take on new meaning and be translated in ‘useful’ ways in the field. Hence, qualification becomes the socio-material process via which new qualities are attributed to measured values in order to become locally malleable *and* stabilised, pre-arranged *and* re-arranged in order to suit local needs. This is the process of attributing new qualities to standardised values that have already been commonly accepted. As this chapter showed, calling one’s data practices as purer than another’s, promoting data collection as a ‘duty of care’ towards communities, assembling different data sources to suit local preferences and needs in a bid to look more democratic and ethical, are all acts of politically (ear-) marking numerical work as a lot more than simply numbers that represent reality as is. Thus, reflexivity becomes a useful instrument in the everyday political struggles that experts fight, not only to collect data ‘values’ from the field, but crucially ‘to establish what value is’ (Graeber, 2001, p. 88).

Therefore, and to conclude, unlike the predominant focus in the literature posing reflexivity as almost ‘hidden’ and happening on the level of the individual (Porter, 2020; Scott, 2008), I showed how reflexivity is mobilised, either through socio-material work (data storytelling and visualisations) or at the micro-level of expert practices (instrumental reflexivity). Thus, reflexivity is a key resource in pushing for datafied governance, especially in a context of increased emphasis on democratisation and decolonisation: reflexivity allows the assignment of political values (values with ‘heart and soul’) back to the measurement of statistical values, in order to enlist participation, facilitate inclusion and thus further enhance quantification as the only available means to know and govern education.

REFERENCES

- Allen, W. L. (2018). Visual brokerage: Communicating data and research through visualisation. *Public Understanding of Science*, 27(8), 906–922.

- Amit-Danhi, E. R. (2021). Facts of (financial) life: Political power and re-visualization strategies. *Convergence*. <https://doi.org/10.1177/13548565211045533>
- Arora, P. (2016). The bottom of the data pyramid: Big data and the global south. *International Journal of Communication*, 10, 1681–1699.
- 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(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., Grek, S., & Tichenor, M. (2023). The rise of the reflexive expert? Epistemic, careful and instrumental reflexivity in global public policy. *Global Social Policy*. <https://doi.org/10.1177/14680181221145382>
- Begkos, C., & Antonopoulou, K. (2020). Measuring the unknown: Evaluative practices and performance indicators for digital platforms. *Accounting, Auditing & Accountability Journal*, 33(3), 588–619.
- Bekkers, V., & Moody, R. (2014). *Visual culture and public policy: Towards a visual polity?* Routledge.
- Benschop, R. (2009). STS on art and the art of STS: An introduction. *Krisis: Journal for Contemporary Philosophy*, 1, 1–5.
- Best, J. (2014). *Governing failure: Provisional expertise and the transformation of global development finance*. Cambridge University Press.
- Bevan, G., & Fasolo, B. (2013). Models of governance of public services: empirical and behavioural analysis of econs and humans. In A. Oliver (Ed.), *Behavioural Public Policy* (pp. 38–62). Cambridge University Press.
- Bevir, M. (2011). *The SAGE handbook of governance*. SAGE.
- Blatchford, I., & Blyth, T. (2019). *The art of innovation: From enlightenment to dark matter*. Penguin.
- Bloor, D. (2013). ‘The strong program’, ‘Sociology of scientific knowledge’, ‘The performative theory of institutions’. In B. Kaldis (Ed.), *Encyclopedia of Philosophy and the Social Sciences* (pp. 961–962, 945–948, 706–708). SAGE Publications.
- Boltanski, L., & Thévenot, L. (2006). *On justification: Economies of worth*. Princeton University Press.
- Borgdorff, H., Peters, P., & Pinch, T. (2020). *Dialogues between artistic research and science and technology studies*. Routledge.
- Bua, A. (2019). Democratic innovations and the policy process. In S. Elstub & O. Escobar (Eds.), *Handbook of democratic innovation and governance*. Edward Elgar Publishing Limited. ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=6005288>

- Callon, M. (2002). Pour En Finir Avec Les Incertitudes? *Sociologie Du Travail*, 44(2), 261–267.
- Callon, M., & Law, J. (2005). On qualculation, agency and otherness. *Environment and Planning D: Society and Space*, 23, 717–733.
- Cochoy, F. (2008). Calculation, qualculation, calculation: shopping cart arithmetics, equipped cognition and the clustered consumer. *Marketing Theory*, 8(1), 15–44.
- Decuypere, M., & Landri, P. (2020). Governing by visual shapes: University rankings, digital education platforms and cosmologies of higher education. *Critical Studies in Education*, article in press
- Desrosières, A. (2015). Retroaction: How indicators feed back onto quantified actors. In R. Rottenburg, S. E. Merry, S. Park, & J. Mugler (Eds.), *The world of indicators* (pp. 329–353). Cambridge University Press.
- Espeland, W., & Yung, V. (2019). Ethical dimensions of quantification. *Social Science Information.*, 58(2), 238–260. <https://doi.org/10.1177/0539018419851045>
- Espeland, N. W., & 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.
- Espeland, W. N., & Lom, S. E. (2015). Noticing numbers. In M. Kornberger, L. Justesen, & A. K. Madsen (Eds.), *Making things valuable* (pp. 18–37). Oxford University Press.
- Espósito, E., & Stark, D. (2019). What’s observed in a rating? Rankings as orientation in the face of uncertainty. *Theory, Culture & Society*, 36(4), 3–26.
- Falisse, J.-B., & McAteer, B. (2021). Visualising policy responses during health emergencies. Learning from the COVID-19 policy trackers. *Convergence*. <https://doi.org/10.1177/13548565211048972>
- Fileborn, B., & Trott, V. (2021). “It ain’t a compliment”: Feminist data visualisation and digital street harassment advocacy. *Convergence*. <https://doi.org/10.1177/13548565211045536>
- Fontdevila, C. (2021). *Global governance as promise-making. Negotiating and monitoring learning goals in the time of the SDGs* (PhD thesis). Facultat de Ciències Polítiques I Sociologia, Universitat Autònoma de Barcelona.
- Gamble, A. (1994). *The free economy and the strong state: The politics of Thatcherism*. Macmillan.
- Gatzweiler, M. K., & Ronzani, M. (2019). Prospective sensemaking and thinking infrastructures in a large-scale humanitarian crisis. *Research in the Sociology of Organizations*, 62, 85–112.
- Graeber, D. (2001). *Toward an anthropological theory of value: The false coin of our own dreams*. Springer.

- 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.
- Grotlüschen, A., & Buddeberg, K. (2020). PIAAC and the South—Is Southering the new othering? Global expansion of dominant discourses on adult literacy. *European Journal for Research on the Education and Learning of Adults*, 11(2), 167–181.
- Hoggett, P., & Thompson, S. (Eds.). (2012). *Politics and emotions: The affective turn in contemporary political studies*. Bloomsbury Publishing.
- Jones, C. A., & Galison, P. (2014). *Picturing science, producing art*. Routledge.
- Jones, R., Pykett, J., & Whitehead, M. (2013). *Changing behaviours: On the rise of the psychological state*. Edward Elgar.
- Kennedy, H., & Hill, R. L. (2018). The feeling of numbers: Emotions in everyday engagements with data and their visualisation. *Sociology*, 52(4), 830–848. <https://doi.org/10.1177/0038038516674675>
- Kennedy, H., Hill, R. L., Aiello, G., et al. (2016). The work that visualisation conventions do. *Information, Communication & Society*, 19(6), 715–735.
- Kornberger, M. (2017). The values of strategy: Valuation practices, rivalry and strategic agency. *Organization Studies*, 38(12), 1753–1773.
- Latour, B. (1986). Visualization and cognition. *Knowledge and Society*, 6(6), 1–40.
- Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of the scientific facts*. Sage Publications.
- Law, J. (2017). STS as Method. In U. Felt, R. Fouché, C. A. Miller, & L. M. Smith-Doerr (Eds.), *The Handbook of Science and Technology Studies* (pp. 31–57). MIT Press.
- Le Grand, J. (2003). Motivation, agency, and public policy. Of knights, knaves, pawns & queens. Oxford University Press.
- Le Grand, J. (2007). *The other invisible hand. Delivering public services through choice and competition*. Princeton University Press.
- Lee, C. W., McQuarrie, M., & Walker, E. T. (Eds.). (2015). *Democratizing inequalities. Dilemmas of the new public participation*. NYU Press.
- Lefsrud, L., Graves, H., & Phillips, N. (2020). “Giant toxic lakes you can see from space”: A theory of multimodal messages and emotion in legitimacy work. *Organization Studies*, 41(8), 1055–1078.
- Merry, S. (2019). The sustainable development goals confront the infrastructure of measurement. *Global Policy*, 10, 146–148.
- Merry, S. E. (2016). *The seductions of quantification: Measuring human rights*. University of Chicago Press.
- Mesch, C. (2013). *Art and politics: A small history of art for social change since 1945*. Bloomsbury.
- Mol, A. (2002). *The body multiple*. Duke University Press.

- Montoya, S. (2017, July 12). A pragmatic and unified approach to measure learning globally (Blog post UIS). Retrieved from: http://uis.unesco.org/en/blog/pragmatic-and-unified-approach-measure-learning-globally?undefined&wbdisable=true*
- Moser, I., & Law, J. (2006). Fluids or flows? Information and qualculation in medical practice. *Information, Technology and People*, 19(1), 55–73.
- Pandey, A. V., Manivannan, A., & Nov, O. (2014). The persuasive power of data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 2211–2220.
- Pollock, N., & D’Adderio, L. (2012). Give me a two-by-two matrix and I will create the market: Rankings, graphic visualisations and sociomateriality. *Accounting, Organizations and Society*, 37(8), 565–586.
- Porter, T. M. (2020). Trust in numbers. In *Trust in numbers*. Princeton University Press.
- Pykett, J., Jupp, E., & Smith, F. M. (2016). Introduction: Governing with feeling. In E. Jupp, J. Pykett, & F. M. Smith (Eds.), *Emotional states: Sites and spaces of affective governance* (pp. 2–18). Taylor & Francis.
- Quattrone, P. (2017). Embracing ambiguity in management controls and decision-making processes: On how to design data visualizations to prompt wise judgement. *Accounting and Business Research*, 47(5), 588–612.
- Reinicke, J. (2015). The politics of values. In M. Kornberger, L. Justesen, J. Mouritsen, & A. K. Madsen (Eds.), *Making things valuable*. Oxford University Press.
- Rose, G., Degen, M., & Melhuish, C. (2014). Networks, interfaces and computer-generated images: Learning from digital visualizations of urban regeneration projects. *Environment & Planning d: Society & Space*, 32, 386–403.
- Rottenburg, R. (2009). *Far-fetched facts: A parable of development aid*. The MIT Press.
- Scott, J. C. (2008). *Seeing like a state*. Yale University Press.
- Skeggs, B. (2014). Values beyond value? Is anything beyond the logic of capital? *The British Journal of Sociology*, 65(1), 1–20.
- Slager, R., & Gond, J.-P. (2020). The politics of reactivity: Ambivalence in corporate responses to corporate social responsibility ratings. *Organization Studies*, 43(1), 59–80.
- 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. <http://www.jstor.org/stable/285080>.
- Strathern, M. (1987). Out of context: The persuasive fictions of anthropology. *Current Anthropology*, 28, 251–281.

- Streeck, W. (2013). *Buying time: The delayed crisis of democratic capitalism*. Verso.
- Wedlin, L. (2006). *Ranking business schools: Forming fields, identities and boundaries in international management education*. Edward Elgar Publishing.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Navigating the Market of Measurement: Data, Quality, and Competition

1 INTRODUCTION

This chapter discusses the rise of the competition over measurement which has been structuring the relationships between IOs. The production of data to support comparative assessment and evaluation is one of IOs' key organisational remits, therefore, they have vested interests in promoting the implementation of their measures over those of others. Consequently, what we observe in the global governance of education is not merely 'governing by numbers', but rather *a navigation of the market of measurement*; this can often lead to conflicts and controversies over statistical data collection, as well as new partnerships and collaborations. Thus, it becomes obvious that it is not merely epistemic authority that governs the production of quantification. Rather, a market logic affects the way data are constructed, collected and compared. In this setting, measures are not merely assessed based on their epistemic qualities—for example, how well they capture the reality of higher education—but rather in their 'market share', i.e. the number of countries and agencies agreeing to participate and contribute to the work of measurement.

This chapter studies two empirical cases of 'measurement markets': the first examines the rise of the 'learning assessment market' that has emerged towards the measurement of learning outcomes for the SDG4. The second case moves away from the transnational space of the SDGs, in order to analyse a case a lot closer to home; that is, the rise of higher

education (HE) in Europe as a field of measurement and competition. Whereas most literature that studies HE focuses on the competition between European universities in terms of rankings and research power, here we focus on a different kind of struggle; namely, this is a story of the interdependence of higher education quality assurance agencies, as they struggle for position and purpose in the dense space of quality assurance in education in the EU. As we will see, since the foundation of the EU, higher education has always been central to Europeanisation, a process that intensified with the Bologna Process (1999). However, as with Europeanisation itself, the work of HE quality assurance has transformed into an organic, living entity, taking root and growing in unexpected ways.

2 THE MARKET OF LEARNING ASSESSMENTS

Previous chapters in this book have already discussed the conflictual character of the negotiations surrounding the focus of the education sustainable development goal (SDG4). However, apart from the competitions and struggles that building the architecture of decision-making towards the SDG4 created, there has also been substantial contestation around the selection of measures for the different indicators, once the latter were decided upon. As Clara Fontdevila eloquently describes, ‘by the mid-2010’s, there were several cross-national assessments (CNAs) in place, but no consolidated methodology to equate and harmonize them’ (2023, p. 6). Although there have been multiple other national datasets and providers that eventually also came into the picture, in this section, we are going to focus specifically on learning assessments, given the significance they acquired post-PISA success, but also, as we will see, the multiplicity and competition over which ones will dominate the ‘market’ (Table 1)

Such diversity of measurements of learning outcomes (especially when contrasted to the lack of data for other indicators such as citizenship or gender equity) reflects the ‘learning turn’ and the emphasis on outputs rather than education inputs, as was described in Chapter 2. What is of interest in this section however is that the technical challenges for the harmonisation of these tools were not hidden, but discussed openly by IO actors in two UNESCO World Education Blogs in 2019. Interestingly, in these two blogs, their authors, Silvia Montoya, head of the UNESCO

Table 1 The market of learning assessments (Fontdevila, 2023)

<i>Overview of cross-national assessments</i>			
<i>Target potentially informed by the assessment</i>	<i>Grade or target population</i>	<i>Name of the assessment</i>	<i>Domain-Literacy (L) or Numeracy (N)</i>
–	1	EGMA	N
	1	EGRA	L
4.1.1a	2	PASEC	L, N
		EGMA	N
	3	EGRA	L
		LLECE	L, N
		EGMA	N
		EGRA	L
	4	<i>PIRLS/ePIRLS</i>	<i>L</i>
		<i>TIMSS</i>	<i>N</i>
		<i>LaNA</i>	<i>L, N</i>
		PILNA	L, N
5		SEA-PLM	L, N
		LLECE	L, N
4.1.1b	6	PASEC	L, N
		SACMEQ	L, N
	8	PILNA	L, N
		<i>LaNA</i>	<i>L, N</i>
4.1.1c	15 y.o.	<i>TIMSS</i>	<i>N</i>
	14–16 y.o.	<i>PISA</i>	<i>L, N</i>
–	–	<i>PISA-D</i>	<i>L, N</i>
		ASER, Uwezo (5–16 years-old)	L, N

KEY EGMA: Early Grade Mathematics Assessment, EGRA: Early Grade Reading Assessment, PASEC: Programme for the Analysis of Education Systems, LLECE: Latin American Laboratory for the Assessment of the Quality of Education, PIRLS: Progress in International Reading Literacy Study, TIMSS: Trends in International Mathematics and Science Study, LaNA: Literacy and Numeracy Assessment, PILNA: Pacific Islands Literacy and Numeracy Assessment, SEA-PLM: Southeast Asia Primary Learning Metrics, SACMEQ: Southern and Eastern Africa Consortium for Monitoring Educational Quality, PISA: Programme for International Student Assessment, PISA-D: PISA for Development, ASER: Annual Status of Education Report, Uwezo: Uganda Learning Assessment.

Notes **In bold**: regional assessments; *In Italics*: international; assessments; Regular font: assessments of foundational skills and population-based assessments.

Includes only those domains relevant for global reporting purposes.

Source Fontdevila's elaboration on the basis of Treviño and Órdenes (2017) and UIS (2016a)

Institute of Statistics, and Luis Crouch, Senior Economist at RTI International, are reflecting on the problems of the ‘learning assessment market’ (Montoya & Crouch, 2019a):

Measuring learning outcomes is key to the Sustainable Development Goal for education (SDG 4). There are about a dozen indicators that measure learning outcomes. Data for these indicators are provided via a market. It may seem odd to think so but think about it for a moment: there are data producers, there are data consumers (countries, policymakers, international agencies and researchers), and there are goods and services exchanged for money (prices) to produce the assessment data. (Montoya & Crouch, 2019a)

Here, two key actors in the making of the SDG4 are open about the ways that learning assessments have become an industry with ‘sellers’ and ‘buyers’, as well as money changing hands in the process. The authors discuss the difficulties of navigating these measures and trying to work with them in order to make them comparable. Interestingly, they reflect not only on the challenge of their own work of commensuration, but also on the ‘market’ itself, which they characterise as a failing one, since it apparently does not adhere to any of the rules that well-functioning markets do:

While the specifics of a market will obviously vary, there are two central questions: does it allocate resources efficiently and equitably? In this blog, we ask this of the learning assessment market, and find the answers fall short.... with learning assessments, there is product differentiation. In fact, no important ‘product’ sold in the learning assessment market is the same as any other, and *organizations purposefully differentiate*. Some assessments are about skills needed for the labour market, others are curriculum-based. Some are designed for primary education, others focus on lower secondary. Some are citizen-led, others are government-led. And so on. (Montoya & Crouch, 2019a)

Although the core idea of the quotation above is that an efficient market would require some uniformity, rather than differentiation, learning assessment producers intentionally differentiate their products so that they appear to offer a tool that is unique and most closely meets the needs of the country ‘buying’ the product. According to Montoya and Crouch, such differentiation goes against the rules of market efficiency.

Nonetheless, in the competition for producing expert knowledge, such competition makes good sense: expert organisations need to differentiate their goods in order to compete in the very dense space of data production for governance. As necessary as alliances and collaborations may be, so is retaining their unique branding and contribution. Indeed, most conflicts between IOs arise when on the one hand they agree to collaborate, while on the other they ‘push’ for their own data instruments and tools, with the World Bank being seen as the usual perpetrator of such moves.

There are also significant barriers for possible competitors to enter the market because it is costly to build a set of good learning assessment questions. New providers typically emerge only to provide a differentiated product. For example, there are assessments serving different geographies (such as initiatives in East Asia) or offering different ways of administering and engaging with the community (citizen-led assessments) as well as different education levels (e.g. the Collegiate Learning Assessment, a higher education standardized test in the United States). (Montoya & Crouch, 2019a)

Another common issue is the differentiation of measures that emerges through the regionalisation of assessments and the difficulties to align them in the production of global data. In order for data to be seen as useful, data producers create assessments that allow countries to compare themselves with their neighbours, rather than with countries at the other side of the globe. Similarly, data producers decide on the focus of the assessment depending on need: as the table above shows, most assessments are focused on the measurement of literacy and numeracy, while fewer ones focus on skills and problem-solving.

There is also price discrimination. Not all countries pay the same. There is some negotiation on price and different levels of subsidies. There is also intermediation. Prices in many cases are negotiated between third party payers (e.g. development partners) and the producer. This can be a good thing in some ways (e.g. the poor pay less) but it also results in non-transparency of prices. (Montoya & Crouch, 2019a)

Finally, this last quotation is a reminder of the costs of producing such learning assessments and the ways these are distributed and differentiated depending on historical and political ties, zones of influence and

donor choices. A common gap in the examination of expert knowledge production is its cost—the cost to produce it and the cost to buy it. As Crouch and Montoya suggest, the lack of transparency around costs leads to further competition, lack of trust and continued high prices, especially for those countries that may struggle most to pay for these data: ‘For example, countries are often led to believe that by joining an international assessment they will benefit from economies of scale. Yet why is it that the fees never seem to go down as the pool of participants grows?’ (Montoya & Crouch, 2019b).

Although the two blogs are only a small snapshot into the world of the ‘learning assessment markets’, the choice of language in both blogs is telling: there are mentions of the need to construct ‘consumer guides’, ‘efficiencies’ and the need to ‘provide more transparent price information’. Crouch and Montoya suggest that ‘the processes whereby consumers and producers interact is a black box’ and thus propose the creation of ‘physical marketplaces’ (Montoya & Crouch, 2019b):

Most of us like touching and feeling things we buy. If we are buying a bicycle or car, it is sensible to try it—even if we end up making the final purchase online. The learning assessment market should offer the same experience—a place where users, producers, and international organizations can meet and make sales pitches. (Montoya & Crouch, 2019b)

The production of global learning data is therefore not produced entirely on epistemic grounds; it is, as we have seen, a matter of political choices over time, as well as a ‘product’ of stark competition in the market of measurement, where data producers have to ‘make sales pitches’ to promote their measures over those of others. However, ‘physical’ this market of measurement can be, there are limitations and visual warnings offered to ‘shoppers’, too (Fig. 1):

Moving away from the transnational space of the production of expertise, the next sections will focus on the case of the quality assurance market in higher education in Europe.



Fig. 1 Image from the World Education Blog ‘The Learning Assessment Market: pointers for countries’ (Montoya & Crouch, 2019b)

3 THE CASE OF THE QUALITY ASSURANCE MARKET IN HE IN EUROPE

The aim of this section is to analyse the growth and complexity of Quality Assurance (QA) in higher education (HE) in Europe, as a way of understanding the multifaceted and continuously developing market of measuring and quality ‘assuring’ universities in Europe. Indeed, the rise of a complex epistemic infrastructure (Tichenor et al., 2022)—with new materialities and actors—has led to the development of intricate webs of education actors and data that have strengthened the emergence of a European education policy space. In fact, the latter is not an imagined space any longer, either to be embraced or resisted; it has become the officially announced and strategically drawn European Education Area,¹ as a single and unified EU policy arena, and thus a strategic area of interest that has to be ‘softly’ governed via a multiplicity of measures and agencies.

None of these developments are of course new. Since the turn of the century, a powerful device for the construction of the European education policy space has been the incessant generation of statistical data to monitor performance (Grek, 2016; Lawn, 2011; Lawn & Grek, 2012).

¹ <https://education.ec.europa.eu>.

The datafication of education policy (Grek et al., 2020) occurred—and partly led to—a fixation on notions of quality assurance and evaluation (Ozga et al., 2011). Indeed, recent decades have seen the notion of ‘quality’ becoming central to attempts to control and develop both public and private institutions, as evident through the proliferation of terms such as ‘quality assurance’, ‘quality enhancement’, audit and ‘quality monitoring’ (Jarvis, 2014). While industrialisation brought the idea of quality assurance to the fore, as the means by which to ensure mass-produced goods could withstand an ‘objective’ quality test against a set of pre-determined criteria, after the 1980s and the rise of New Public Management, ‘quality’ acquired a double meaning. It now relates not only to the quality of products or services but also, crucially, represents a key criterion for judging how organisations are run. ‘Quality gurus’ emerge and quality assurance processes travel from organisation to organisation (Power, 2003). Quality must be measured quantitatively and at all times, and it represents the means through which organisations can be compared and become ‘known’ to citizens/consumers. In the case of transnational policy spaces and political projects, like the EU, quality and all its associated measurement processes, such as those of ‘quality assurance’, become a main mode of ‘soft’ governance (Lawn, 2011), operating through the setting of common benchmarks and standards and the promotion of constant self-regulation as a way to learn and to align oneself with international ‘best practice’.

Since the 1990s, this ‘soft governance’ turn has led to the creation and expansion of a European-level quality assurance market in higher education in Europe (Gornitzka & Stensaker, 2014). QA is often imagined as an instrument for greater internal mobility in Europe, while also advertising and guaranteeing the quality of European skilled labour and knowledge products, in line with European Union goals related to becoming the world’s most advanced knowledge economy. In the following sections, I examine shifts in HE quality assurance, in the form of standards, data and reports. Second, I explore the market of quality assurance actors involved in European QA and measurement processes, their interdependencies and their contestations; I examine European actors, such as ENQA and EQAR, but also the influence of global ones, such as the OECD. Finally, I reflect upon what these explorations reveal about how a market of QA measurement in Europe has evolved over the last two decades, what the position of the Bologna Process has been in these dynamics, and how QA has become a central feature and driver not only

of the Europeanisation of HE per se, but also of the construction of a market of measurement as a whole.

3.1 *Europeanisation as a Concept and as a Research Conundrum*

As discussed extensively elsewhere (Lawn & Grek, 2012), a focus on QA, alongside the expansion in data production and use, and its capacity to flow across Europe (and beyond), illustrates a shift from merely using data to provide a ‘state optic for governing’ (Scott, 1998) into the fabrication of European education as a legible, governable policy space. In *Europeanising Education*, Martin Lawn and I describe the ways that the positioning of policy actors as ‘policy brokers’, that is people who are located in some sense at the interface between the national and the European, ‘translate’ the meaning of national data into policy terms in the European arena, while at the same time continuously interpret European developments in the national space. Adopting the term ‘brokers’ here, I do not intend to paint a picture of national–transnational exchanges, in which policy brokers operate as frontier guards, and members of European organisations act as carriers of a European policy agenda. Instead, I understand Europe to be fluid and changing, and itself swept by international pressures, simultaneously located in and produced by the global, the idea of the European and the national. In order to capture this constantly moving, liquid and undefined European education space, we start the analysis from a slightly more stable ground: its past. Education policy activity in the European Union (EU) could historically be classified in several ways; for example, the Treaty of Rome (1957), the Single Act (1987) and the Maastricht (1992) and Amsterdam (1997) and Lisbon (2009) Treaties could be seen as five stages (1957–1987; 1987–1992, 1992–1997, 1997–2009 and 2009-) (Ollikainen, 1999; Shaw, 1999). The European Education Policy Space was not determined merely by the fairly stable geographical boundaries of a common market: as early as the 1960s, it became a shared project and a space of meaning, constructed around common cultural and educational values. Indeed, from the 1960s to 1970s, the discourse of a common culture and shared histories was slowly being produced as a cluster of facts and myths about the European ‘imagined community’ rising from the ashes of a destructive Second World War. Education policymaking for the ‘people’s Europe’ took the forms of cultural cooperation, student mobility, harmonisation of qualification systems and vocational training (European Commission, 2006).

It did not constitute a purely discursive construct, adding to the list of European myths. It was concretised and pursued through Community programmes, such as COMETT and ERASMUS, involving large numbers of people and travelling ideas (European Commission, 2006). Its impact was arguably limited in relation to the ways European education systems constructed their curricula and tools of governance; subsidiarity was the rule. However, regardless of its relatively limited effects, the project of a ‘people’s Europe’ had a clear ambition: to create a distinct European identity and culture—and to use these resources to enable the governing of a shared cultural and political space.

This brief reminder of the foundational characteristics of Europeanisation is important here for two reasons: first, it helps to throw into relief the defining events that turned the European education space from a rather idealistic project of cultural cohesion to a much sharper competitive reality; and second, it enables us to understand how, when and why the discourse of QA entered this space, and with what impact. For example, research reveals the many points of origin identified by national policy actors in relation to policy requirements that demand data collection—these may originate in Europe or from the wider world of OECD, the United Nations Educational, Scientific and Cultural Organization (UNESCO) or the World Bank. Indeed, for the most part, the source of pressures and requirements does not seem to be of great concern. Instead, policy actors focus on ensuring successful outcomes, on producing ‘world-best’ education through the production and use of data: securing competitive performance is the language of high quality and standards. In the aftermath of a global pandemic, and the ‘protectionist’, primarily national, policy responses that it ensued, there are even greater difficulties in identifying a distinctive European Education Policy Space, as policy actors interpret their brokering as a fusion of European and global influences that places pressure on systems to demonstrate success in terms of measurable outcomes. Such developments suggest that the ‘Europe’ of a collective project of shared trajectories, values and aspirations is less visible than in the past, and focuses attention on the kind of space of governance that the growth of data flows in Europe gives rise to. Looked at in this way, we can see that the governing project of a ‘people’s Europe’ is slowly being turned to a project of individualisation—the production of a Europe of individuals, striving to accomplish the next set of goals, indicators and benchmarks. This project is made possible by the existence of networks through which data may flow, and as I will show,

through the competition of a range of measures and monitoring tools that connect individual student performance to national and transnational indicators of performance. Furthermore, the use of these particular technologies of governing signals a shift from the attempted *fabrication* of Europe through shared narratives and projects to its *projection*. By this I mean a shift from the production of Europe through the recording and transmission of its existing characteristics and capacities to the moulding of the future through QA processes that shape and project the individual and the nation forward into lifelong engagement with Europe as the most competitive knowledge economy in the world. It is within this conceptual context that we will turn to an examination of higher education and the impact of QA measures in shaping the field for at least two decades now.

3.2 *Quality Assurance in Higher Education in Europe*

Although the story of efforts for the convergence of higher education in Europe goes as far back as the inception of the European political project in the early 1970s, it was the Bologna Declaration of 1999 that instituted a process that fundamentally reshaped European higher education (Curaj et al., 2018; Enders & Westerheijden, 2014; Schriewer, 2009). While the precise objectives have evolved over time in connection with the work of the Bologna Follow-Up Group (BFUG) and, in particular, the Ministerial Conferences of members, the main goals of the process have concentrated on mobility between, and the compatibility of, higher education systems and the pursuit of quality in higher education (Bergan, 2019). In practical terms, the drive towards these objectives has included a focus on the structuring of systems in accordance with the three-cycle approach (Bachelor, Master's, Doctorate); the creation of an EHEA Qualifications Framework; and the development of common standards and processes for QA (Bergan & Deca, 2018; Brøgger, 2019). This drive resulted in the announcement of an education space of enhanced mobility and competitiveness, the European Higher Education Area (EHEA) in 2010. Extending beyond the borders of the European Union, the EHEA's 49 country members—joined by the European Commission and a range of stakeholder organisations—have all agreed to pursue the goals of the Bologna Process, altering their HE systems to facilitate the mobility of students and staff between EHEA members and to enhance the employability of graduates (Barrett, 2017).

These processes of reform have been accompanied by the creation of a wealth of academic and practitioner publications describing the evolution of Bologna and the EHEA, evaluating the strengths and weaknesses of the EHEA, and prescribing future directions for development. The mammoth edited volumes on higher education within the EHEA by Curaj et al. (2012, 2015, 2018) are a clear example of this body of literature. However, gaining analytical purchase on the transformations within HE since the initiation of the Bologna Process, requires stepping outside of an ‘insider’s perspective’ (Dale, 2007) and viewing the developments in their historical and political context. Corbett (2012), for example, highlights how European higher education cooperation and governance have changed with the onset of the Bologna process and the creation of the EHEA, with new European policy arenas being created where there had been relatively little European-level action. As both Dale (2007) and Corbett (2011) indicate, all this reflects wider transformations in the role of the university in the era of knowledge economies and, in particular, the notion of a Europe of Knowledge (Corbett, 2012; Dale, 2007). The rapid adoption of the push for Bologna reforms and the EHEA—with 45 countries involved by 2005 (Bergan, 2019)—speaks to their political resonance for this changing context.

One of the most significant dimensions of change associated with the Bologna Process has been the role and influence of the European Union in HE. The European Commission’s scope of action in education is restricted by the subsidiarity principle in education, but the Bologna Process has provided a means for the Commission to fulfil its supporting obligations and to work around such limitations (Brøgger, 2016; Capano & Piattoni, 2011). Despite being initially positioned outside the Bologna Process and the development of the EHEA, the European Commission has come to occupy a central role in driving the agenda (Dakowska, 2019; Robertson, 2008). Keeling describes, for example, how the European Commission began to dominate the higher education discourse in the 2000s, with the Commission’s involvement in the language politics around research policy and the Bologna Process contributing significantly to ‘the development of a widening pool of “common sense” understandings, roughly coherent lines of argument and “self-evident” statements of meaning about higher education in Europe’ (2006, p. 209). As Magalhães et al. (2012) explain, part of this process of European consolidation has been the ability of the European Commission to bring together, or to articulate (Veiga, 2019), multiple agendas and

discourses in ways that expand the legitimacy of European-level action in the EHEA and in higher education more broadly. Of particular importance was the drawing together of the development of the EHEA with the economic agenda of the Lisbon Strategy, which sought to ‘to make the Union the most competitive and dynamic knowledge economy in the world’ (Krejsler et al., 2012), and the Modernisation Agenda for universities, inspired by the New Public Management school of thinking (Enders & Westerheijden, 2014). Crucial here also is the ability of the Commission to allocate funding to support activities that align with its conception of what the EHEA should be, especially given the lack of overall EHEA funding (Bergan, 2019). As I will show, we observe an organic growth of actors and datasets in the field of quality assurance in Europe, boosted by the growth and expansion of quantification in policy-making; in this context, quality assurance does not represent merely a tool of governing higher education but has also become a market of measurement, with universities being proclaimed as carriers of ‘global Europe’ and the ‘European way of life’ (EC, 2022).

3.3 *Quality, Data and Standards*

This section explores major shifts that have occurred in the production of data for quality assurance and measurement in European higher education since the inception of the Bologna Process. I focus principally on three key documents and datasets—the Standards and Guidelines for Quality Assurance in the EHEA (ESG) (developed by ENQA, see below), the European Quality Assurance Register for Higher Education (EQAR), and the European Tertiary Education Register (ETER)—which together illustrate the increase in both the scope and the complexity of the changing QA architecture. Crucially, these three developments have helped create the foundations, and interconnections, that facilitate further diversification and expansion of the QA market of measures.

One of the first organisations to emerge in connection with the initial Bologna developments was the European Association for Quality Assurance in Higher Education (ENQA). ENQA is a stakeholder organisation whose membership comprises principally of quality assurance agencies (QAAs). QAAs perform reviews of higher education institutions and programmes, making them key actors in higher education systems. In addition to serving as the main representative of this key constituency

in European higher education, ENQA has also taken a lead on developing the underlying infrastructure of QA in Europe (Ala-Vähälä & Saarinen, 2009). In 2003, the Bologna Ministerial Communiqué called on ENQA alongside the European Students' Union (ESU, previously ESIB), European Universities Association (EUA) and European Association of Institutions in Higher Education (EURASHE)—the other members of what came to be called the E4—to develop an agreed set of standards, procedures and guidelines on QA (E4, 2011). This followed a recognition in the Berlin Communiqué of the Bologna Process that the 'quality of higher education has proven to be at the heart of the setting up of a European Higher Education Area', with Ministers stressing the 'need to develop mutually shared criteria and methodologies on quality assurance' (p. 3).

The outcome of the ENQA-led process was the 2005 creation of the European Standards and Guidelines (ESG), which were adopted as part of the Bologna Process's Bergen Communiqué and which were framed as a step towards greater consistency in QA across the EHEA and enhanced trust and qualification recognition between different contexts. The ESG outline standards and guidelines for different types of QA processes and the different actors involved in them. The standards set out broad and basic requirements in order for institutions and QAAs to be compliant with the ESG, such as that 'institutions should have formal mechanisms for the approval, periodic review and monitoring of their programmes and awards' (Standard 1.2). The guidelines provide 'additional information about good practice and in some cases explain in more detail the meaning and importance of the standards', although it was not 'considered appropriate to include detailed "procedures"' (p. 11) in the guidelines. The first part of the ESG focuses on internal QA processes within higher education institutions, the second on QA by external actors (i.e. QAAs) and the third on QAAs themselves. For external QA processes, for example, ESG compliance requires that 'Any formal decisions made as a result of an external quality assurance activity should be based on explicit published criteria that are applied consistently' (Standard 2.3), while for QAAs it is required, for instance, that 'Agencies should have clear and explicit goals and objectives for their work, contained in a publicly available statement' (Standard 3.5). Backed by the force of their collective acceptance by the Bologna Process members, these standards and guidelines make claims about how we can come to know the presence or absence of 'quality'.

As suggested by the lack of specification for the ‘mechanisms’, ‘criteria’ or ‘goals and objectives’ mentioned in the three standards above, a key characteristic of the ESG is the openness and ambiguity of the standards and guidelines (Brogger & Madsen, 2021; Gornitzka & Stensaker, 2014). In part, this appears to be a response to the tension present throughout European-level education initiatives between the drive to harmonise practices to facilitate integration, and mobility, and the political and practical realities of Europe’s varied set of education systems. Therefore, while the ESG are working towards the ‘establishment of a widely shared set of underpinning values, expectations and good practice in relation to quality and its assurance’, the report states that diversity and variety are ‘generally acknowledged as being one of the glories of Europe’ and correspondingly ‘sets its face against a narrow, prescriptive and highly formulated approach to standards’. Keeping a studied ambiguity in the formulation of the ESG likely serves as a means of ensuring its acceptability to a wider range of European states and education systems. Rather than a strict standardisation, this might be seen as ‘setting the outer borders within which there is scope for diversity’, as one actor in the space put it for the Bologna framework at large (SB int.). Such a description of the role and function of the ESG fits particularly well with the conceptualisation of this space as an epistemic infrastructure; that is, building the conditions and structures that, at a later stage and possibly by other actors, can be ‘filled’ with new inscriptions and procedures that will make the infrastructure intelligible and useful and grow it anew. Indeed, as another interviewee articulated, the balancing act of the ESG has been to have it ‘prescriptive enough in order to induce the change needed, but also general enough to have so many countries being able to work with it’ (CG int.). Perhaps because of this breadth and ambiguity, the ESG has been one of the most successful harmonising elements of the Bologna Process (Bergan, 2019). Pointing to the transformative power of the ESG, the same interviewee commented, for instance, that QAAs can push reforms with governments by saying that ‘we have the standards, and we have all the colleagues in Europe that are doing it like this, and then we have to align’ (CG int.); peer pressure is therefore strong, one of the most influential qualities of governing by data.

In 2015, the ESG were updated to reflect changes that had occurred with respect to other elements of Bologna, such as qualifications frameworks, as well as broader shifts, for example towards student-centred learning (ESG, 2015). While compliance is by no means universal, the

initial success of the ESG encouraged this evolution and expansion in scope. The inclusion of a new item in the ESGs, or indeed a changed interpretation, likely gives a higher likelihood of members adjusting their systems to incorporate the new directions. As one interviewee put it, ‘people think that if they put something more in the ESGs that it has the chance to be really implemented tomorrow in those... 49 countries which are members’ (CG int.). In recent years, there have been plenty of prompts for further amendments to the ESG in connection with the popularisation of micro-credentials, for example, and the spread of digital learning approaches associated with the COVID-19 pandemic and other longer-term trends. While initially presented as a simple, technical instrument for QA practices, the ESG can be seen here to act as a governing instrument in higher education (Stensaker et al., 2010), with the potential for alterations to and expansions of the material infrastructure of the ESG to reflect new strategic choices and policy trends and, crucially, to induce corresponding changes in the education systems of member countries.

Further, the foundational nature of the ESG can be seen in the case of the European Quality Assurance Register (EQAR), which was created in 2008 as part of following up on one of the recommendations of the initial ESG report. EQAR was the first legal entity to be created through the Bologna Process, and it functions, in some ways, as a guardian of the ESG. QAAs apply to be part of the register, and thus legitimated as trustworthy agents, and are only listed if they are judged to be compliant with the ESG by EQAR’s Register Committee (EQAR, 2020). Through this process, EQAR transforms ‘QA agencies in Europe into QA agencies of Europe’ (Hartmann, 2017, p. 319). Register decisions are made on the basis of external reviews of QAAs that are generally coordinated by ENQA, who, along with the rest of the E4, are founding members of EQAR. The existence and effective functioning of EQAR and, to some extent, ENQA depends, therefore, on the ESG. Part of the power of EQAR and ENQA, however, is their ability, emerging from their recognised responsibility for carrying out the above duties, to create procedures and systems for interpreting the ESG so as to decide on compliance. The way in which these internal procedures and systems operate has the potential to affect which QAAs are labelled as EQAR registered, with acceptance on the register opening doors to performing QA activities in different countries, as well as which higher education institutions and programmes are recognised as being vetted by an EQAR-registered agency, which can influence,

for example, how qualifications are recognised (or not) as students and graduates move between contexts.

Through its register processes and reporting procedures, EQAR has been a key driving force behind an impressive expansion in the epistemic infrastructure around QA in Europe. As well as the reports prepared for admission onto the register and periodic renewal, EQAR also requires reports whenever a registered agency adjusts their practices in a way that might have an impact on their compliance with the ESG. A major expansion in EQAR's data flows and capabilities came in 2017 when EQAR launched the Database of External Quality Assurance Results (DEQAR). DEQAR collects and collates data not just on the QAAs that are part of the register but also, through the reports submitted by those QAAs, on the institutions and programmes that those QAAs have reviewed (EQAR, 2021). As of June 1st, 2022, DEQAR contained nearly 74,000 reports on over 3000 institutions. The foundational structure of the ESG is, again, key here, as one interviewee described: 'DEQAR of course is also very closely related to the ESG standards for higher education and I think you couldn't expand it to another sector or copy it or replicate it into another sector without having a similar kind of agreed European standard available.... If you don't have an agreed standard, then what is the meaning of being in a database, what does it stand for?' (CT int.). The processes of harmonisation connected with the ESG, therefore, have allowed for data produced across countries, QAAs and HE institutions to be transformed into European data and metrics.

The market relating to QA does not exist in isolation but is interlinked with other infrastructures and projects. Examining a third key development, the creation and growth of the European Tertiary Education Register (ETER), provides an example of such interlinkages and helps illustrate the increasing complexity of the market of quality measures in European higher education. ETER started as an academic project funded by the European Commission, which has also supported EQAR and ENQA. ETER sought to respond to an absence in the higher education data infrastructure in Europe, as one interviewee put it: 'a core function of ETER is to provide a list of institutions. You might think it's a stupid task, but such a list did not exist before ETER in Europe' (BL int.). Although conceptually simple, the creation of the register requires important processes of categorisation, standardisation and commensuration, which have built on existing data standards while agreeing and deciding upon new ones. The significance of simply having such a register available

should not be understated, with an underlying standardised way of recognising and recording institutions and their characteristics being extremely valuable for the potential interoperability of different higher education data systems in Europe. Crucially, the existence of such a dataset opens the door for more extensive analysis of the state of European education through the use of the student, graduate, financial and other data collected for each institution in the database.

Through the DEQAR Connect project, funded by the European Commission, the quality assurance and measurement infrastructure provided by ETER and DEQAR have been linked together. As an interviewee described, DEQAR uses ‘ETER as an underlying data source of basic institutional information’, noting further that EQAR ‘only added the quality assurance related information to it’ (CT int.). Working in combination with ETER’s infrastructure on institutions allows EQAR to now present information on, for example, the proportion of a country’s students that are studying at an institution that has been reviewed by an EQAR-registered agency. This represents a significant expansion of the data that EQAR can provide and also moves EQAR closer to dealing with higher education institutions rather than just QAAs. Furthermore, in addition to being connected with ETER, DEQAR data is now being integrated into the workflows of national recognition centres, which offer authoritative advice and guidance to higher education institutions on the recognition of qualifications and assessments (CT int.). This points to an important connecting together of national-level infrastructure associated with qualification recognition with European-level infrastructure concerning QA.

As well as being significant in their own right, these three developments are illustrative of the broader proliferation of a European-level development of the data infrastructure that supports the construction of a ‘market’ that measures quality in higher education (Gornitzka & Stensaker, 2014). Other spaces for discussions on QA have also been built, principally the European Quality Assurance Forum, which generate materialities in the form of reports, minutes, presentations and more. Furthermore, in 2012, Eurydice took charge of the Bologna implementation reports, as they came to be called; the latter have become a central vehicle for evaluating movement across EHEA countries on the core commitments of Bologna, including on QA and recognition. Significantly, Eurydice draws on data and insights from a range of actors in European higher education in order to compile the reports, including EQAR and

ENQA, pointing to the significance of the interlinkages between actors, the second layer of this market of measurement, to which we will turn next.

3.4 *The Market of Actors*

No market could have expanded to the extent and complexity that quality assurance in European universities has over the last 20 years without the efforts of a range of key actors. Following on from the previous discussion, this section focuses on the ways the market of measurement in HE quality assurance has extended to include a range of organisations that are creating new interdependencies and alliances but also new conflicts over policy influence and direction.

One of the more established actors in the field is the aforementioned ENQA. It has seen its influence grow substantially during the last decade, moving from being one of the many stakeholders in the Bologna Process to a much more strategic and policy-oriented role. ENQA was established in 2000 as the European Network for Quality Assurance in Higher Education, only to be renamed four years later to as an ‘Association’ (Ala-Vähälä & Saarinen, 2009). Although its remit from its inception has been to ‘represent QAAs in the EHEA’, to ‘support them nationally’ and to ‘provide them with services and networking’, in recent years its influence has grown. While it is primarily a stakeholder organisation, ENQA has developed a significant role in driving policy concerning QA and is trying to steer the field in new ways (Sarakinioti & Philippou, 2020). ENQA has played a key role in creating, updating and disseminating the ESG, as explained previously. However, according to its strategic plan 2021–2025, ENQA is also pursuing ‘knowledge-based development’ and exploring ‘new ways of quality assurance’ by becoming a forum for ‘...facilitating the discussion on any changes in higher education and its provision and the consequences these changes may entail’. (ENQA, 2021).

Such a broad strategic vision in terms of shaping the field has become a significant aspect of ENQA’s work. ENQA sees its role as a policy actor, strategically placed in close proximity to the European Commission: ‘ENQA is based in Brussels for a reason. So it’s mostly the director that is based there, who is joining different types of activities, meetings with the Commission’ (CG int.). ENQA derives its status from, on the one hand, its established connection with the BFUG through being a

consultative member and, on the other, the sheer strength of the number of organisations it represents:

The weight of ENQA is being given by the members. So when you go to a table, when it's about higher education policy, and then there you represent 55 quality assurance agency members which are compliant with the ESGs from 40 countries, and then you also represent 55 affiliates also from outside Europe, that also makes you an important network. (CG int.)

Indeed, the expansion of the work and of ENQA's influence to other world regions has given it particular momentum. Not only does this increase networking, but also, crucially, it promotes the standing of European higher education as a global higher education actor:

There are other networks of quality assurance agencies from all over the globe, African, Asian, United States, so the collaboration with those networks is important. So what we try to do is to learn from each other, but also our objective is to promote the European standards and guidelines, because of course we believe they are good. (CG int.)

A particularly revealing example of complex interdependencies and contestations in the QA market is the relationship of ENQA with its sister organisation, EQAR. Throughout our examination of the two organisations, there has often been the potential for confusion—not only by us as researchers but crucially also in the field itself—about the distinctions between the two organisations and their work (Huisman et al., 2012). This seems to spark an inclination to differentiate one's own organisation from the other as a way of sustaining the need for the continued existence of both, especially in a field ridden by complaints for duplicity of efforts and over-reliance on bureaucratic form-filling:

EQAR is not the one that is developing the policies or providing services to the members or representing them. They are just a register.... Of course, if for example there is a discussion on revising the ESGs of course they will be involved. But maybe you know that EQAR was founded by the E4. So ENQA is the founding member let's say of EQAR. So they are our kids in a way. (CG int.)

EQAR actors, however, do not necessarily see themselves as 'just a register'. They are also an organisation whose role has evolved and grown,

such that EQAR can now suggest that they can and should influence policy on QA in more fundamental ways:

I would say that the role evolved over this, well, now nearly 15 years in two ways. On the one hand, let's say, from the very beginning EQAR was a very technical and bureaucratic organisation....But then I think very soon EQAR also became involved as an organisation that, let's say, informs the policymaking discussions in the Bologna Process, because of course the governments and other stakeholders were keen to also have EQAR there as an organisation that can give some input and share the expertise that we make and gather from this work of registering agencies, of reviewing which agencies comply with the ESG and so on. And that has become or that has grown little by little over the years and now also there is quite some work done on maintaining a knowledge base on our website, on analysing what is happening in quality assurance in Europe. (CT int.)

Perhaps more so than the micro-disagreements of how the hierarchy or the dependencies among these organisations work, or the extent to which there is a degree of mission overlap or not, what is interesting here is that the growth and expansion of the market of measures (and the agencies that produce them) is seen as 'organic'. For EQAR, an example of this spiralling of work into different directions and branches is the establishment of DEQAR. On the one hand, DEQAR was described as an 'obvious' or 'not that far-fetched idea'. On the other, however, it has been portrayed as 'a major change of our role in these 13 years...[since] ...now we are dealing with the level of higher education institutions by having a database of them and that's of course quite a big difference for our work' (CT int.). In other words, although EQAR's primary role was to work with QAAs, the expansion thanks to the creation of DEQAR means that EQAR now has links not only with QAAs but also with European universities themselves. Such an organic growth and expansion of the QA activities is extraordinary and well beyond what the Bologna Process set out to achieve. ENQA and EQAR perform a lot more than just the technical role of inspecting HE institutions on the basis of the ESGs. Sitting at the BFUG table as experts in QA and representatives of QAAs, they contribute to shaping the future and strategic direction of EHEA. Furthermore, the market of measurement remains intact since, while maintaining their networking function and their allyship with other QA organisations, they continue to preserve their own unique contribution and presence in the field.

A second key actor in the broader field of measuring and evaluating quality in European Higher Education is the Organisation for Economic Cooperation and Development (OECD). Although the OECD is best known in the field of education for the establishment and success of the Programme for International Student Assessment (PISA), less known but equally significant is their work in other areas and especially in higher education. Examples of this work in the European space abound. The Labour Market Relevance and Outcomes of Higher Education (OECD 2022) is one such project, and it receives substantial support from the European Commission, with the main participant nations being Austria, Hungary, Portugal and Slovenia. The project explores issues such as the emergence of ‘alternative credentials’ and the use of ‘big data’ to understand graduate skills and digitalisation in higher education. Previous project participants were Norway, Mexico and four US states. The global nature and reach of the OECD is a valuable resource in the efforts to establish the EHEA as a global player. The ability of the OECD to offer comparative data from other competitor world regions is one of the main reasons for its increasing involvement with issues of quality in HE. In its work to extend the comparisons and the evidence base beyond Europe, the OECD has made use of connections with the ETER project, as explained by an OECD interviewee:

We’ve been involved in that process for, I think, at least on the advisory board since the project’s inception... And part of what we’re doing at the moment is trying to develop the similar data source that draws on ETER, but also draws on other national data collections that are available for the non-EU/OECD countries.

Establishing such international comparisons and linking the European HE quality processes with those of other OECD countries is a key endeavour for both the European Commission and the OECD (Grek, 2014, 2016; Sorensen & Robertson, 2020). Correspondingly, this is a well-established collaborative relationship that has grown substantially over the years, as another interviewee shares: ‘We’re involved, they invite us to all their working meetings and likewise we invite them to ours (DC int.)’. Of course, it is not only the Commission that benefits from the OECD’s expertise. This is a two-way relationship that influences and is

advantageous to both. The OECD benefits from the data the Commission generates as well as, crucially, from the funding available from the Commission for their work:

We actually use quite heavily the EU's surveys, for example, the Labour Force Survey and, similarly, labour force surveys in the non-EU countries to assess a range of labour market outcomes of higher education.... So we run this policy survey and when we're running that, we would obviously take into account work that the European Commission has done previously. We would consult with them on that to make sure that we're not, first of all, duplicating the work, because obviously we all have to be efficient, and then also to make sure that what we're producing makes sense from their perspective as well as from the perspective of our member countries. (GG INT)

In terms of the kind of collaboration and work that the OECD offers, some interviewees stressed the OECD's independence and expert function, as compared with a more politicised Commission, while others emphasised the benefits of continued dialogue, with the Commission setting the strategic direction and the funding and the OECD responding to these policy priorities. Similar to other areas of education policy, the relationship between the OECD and the European Commission is a 'symbiotic relationship', reaping substantial benefits for both organisations:

The European Union is definitely a voice in our, you know, through the European countries that sit in the Group of National Experts. Certainly we would understand very well the priorities of the European Union. (GG int)

'There has evolved a division of labour and a symbiotic relationship between the Commission and the OECD over the last 15 years or so...They're the people with the wallet! So, in a sense, we're more likely to be working within the framework of problems and priorities that they've identified, so it might be the case that the Commission will say to us, we're really concerned about digitalisation in higher education, in which case we would say, oh, well, we agree, we think that's a really important topic and we could support you in a couple of different ways, here, we'll give you a couple of examples of what we might do. But there, you see how that it's a dialogue' (TW).

Finally, as the above suggests, the European Commission is a powerful actor in European quality assurance and evaluation processes. Through its membership of the BFUG, but more importantly, through its provision of funding and its convening power (Brøgger, 2019; Cone & Brøgger, 2020; Dakowska, 2020), the Commission has been able to influence the education policy direction as a whole, both inside and outside Bologna, and, thus, has often been the driving force behind the building of QA infrastructure. The way in which the European Commission coordinates the higher education space is subtle, yet, over time, it is effective in generating change in actors. The ‘pull’ of the Commission—through its funding, networks, data and indicators, and dominant discourses—changes the field in which European higher education actors operate such that it is that bit more likely that their next step will be in the preferred policy direction of the Commission, resulting, with enough time, in substantial movement in that direction. Note that this does not suggest that the Commission drags other actors along, that actors cannot or do not move away from the Commission’s preferred policy direction, or that the Commission is entirely alone in trying to stack the odds in its favour. Instead, it accounts for how—by making it that bit easier to move with and towards the Commission, due in no small part to it being the ‘wallet’ in the field—the Commission can softly direct the evolution over time of the infrastructure around QA and measurement in European higher education.

To further its ambitions into the future, major developments are being planned by the Commission for continuing the evolution of QA market of measures, including steps towards its regulation and management. The 2022 Communique, for instance, proposes the creation of a European Quality Assurance and Recognition System. This system will build a European space ‘where the quality of qualification is assured, the qualifications are digitised and recognised automatically across Europe, doing away with the bureaucracy that hinders mobility, access to further learning and training or entering the labour market’. It seems that, yet again, the discourse and practices of quality assurance are being put to work for the fabrication of the ideal common Europe, where universities do not just promote a ‘European way of life’ but also bring it to fruition. The practices around quality assurance and evaluation, therefore, do not simply represent technical processes by which mobility in Europe is facilitated. Instead, quality becomes a central governing device in an expansive and ever-changing data infrastructure, through which new strategic directions

are drawn, new and old actors are interlinked and the construction of a market of measurement continues apace.

4 DISCUSSION

Building on a rich set of documents and interview data, this chapter focused on an analysis of how the production of expert knowledge for policy has developed a functional and ever-expanding market of measurement over the last two decades.

A first prominent characteristic of the market of measurement is the ‘organic’ growth of its data and processes, emerging as they did from opportunities perceived and seized by particular actors at particular moments, rather than from being clearly tied to a pre-planned strategic progression. In both the cases of the learning assessment markets and of quality assurance in higher education in Europe, we observe the development of a market of measures through the balancing out of supply and demand, as well as through the initiatives of new actors, as they saw opportunity in the field, both in terms of real returns and—perhaps primarily—as ways to establish an influential position in a rapidly growing field. While the creation of DEQAR and its linking together with ETER, for example, certainly fit within the broad strategic vision of the Bologna Process, they came about because actors *sought to* make use of their assets, i.e. the volumes of reports they had processed and the collaborations they had invested in. This finding chimes with literature that suggests the power of numbers to ‘acquire a life of their own’ (Fourcade, 2016).

Similarly, this organic growth of the market of measurement also points to the role of its multi-layered temporality: on the one hand, the foundational nature of learning assessments as the essential building blocks of any global learning data illustrates the potential for market development to have a *sequential* temporality, where the completion of one block allows for the building of the next. On the other hand, however, the regular, cyclical collection of national and international assessment data speaks to different rhythms of market change and operation; both sequential and cyclical change are operating simultaneously like clock cogs. To explain further, while part of the market of measurement operates having a future orientation, others adopt cyclical and repetitive process that establish a ritual and a way of folding new data into the infrastructure. It is through analysing these events as part of a dynamic market of measurement as opposed to discrete policy events that these temporalities become visible.

In addition, the market of measurement creates competition and differentiation between actors. On the one hand, I examined the relationship between EQAR and ENQA, for example, and their struggles over the relative positions of the two in the QA market and the policy space that it helps to constitute. On the other hand, I also discussed how data producers do not construct uniform assessments, but prefer to differentiate their products geographically or substantively. Over time this has led to attempts at enhanced ‘market brand’ differentiation between data producers and IOs in order to more clearly delineate their respective positions in the market of measurement. One important aspect of the changing roles and positions of actors in this space is the new forms of expertise that the establishment and maintenance of these interlinkages and competitions require. The experts working in this field are no longer merely statisticians and data scientists. Increasingly, as we saw in this and previous chapters, what is needed is a new type of expert: expert brokers who can produce ‘sales pitches’ and can persuade that *their* measurement product is better than the next one. These insights, namely on organic growth, temporality and competition and differentiation, paint a picture of a fluid and significantly multi-polar space of expert knowledge production—an enhanced market of measurement that has, by now, grown too large to fail.

REFERENCES

- Ala-Vähälä, T., & Saarinen, T. (2009). Building European-level quality assurance structures: Views from within ENQA. *Quality in Higher Education*, 15(2), 89–103.
- Bandola-Gill, J., Grek, S., & Tichenor, M. (2022) *Governing the sustainable development goals: Quantification in global public policy*, Palgrave/Springer.
- Barrett, B. (2017). *Globalization and change in higher education: The political economy of policy reform in Europe*. Palgrave Macmillan
- Bergan, S. (2019). The European higher education area: A road to the future or at way’s end? *Tuning Journal for Higher Education*, 6(2), 23–49.
- Bergan, S., & Deca, L. (2018). Twenty years of Bologna and a decade of EHEA: What is next? In *European higher education area: The impact of past and future policies* (pp. 295–319). Springer, Cham.
- Bologna. (1999). Bologna Declaration.
- Brøgger, K. (2016). The rule of mimetic desire in higher education: Governing through naming, shaming and faming. *British Journal of Sociology of Education*, 37(1), 72–91.

- Brøgger, K. (2019). *Governing through standards: The faceless masters of higher education. The Bologna process, the EU and the open method of coordination*. Springer.
- Brøgger, K., & Madsen, M. (2021). An affirmative-diffractive re-reading of the policy instrumentation approach through agential realism and the accreditation instrument. *Journal of Education Policy*, 1–19, <https://doi.org/10.1080/02680939.2021.1938239>
- Capano, G., & Piattoni, S. (2011). From Bologna to Lisbon: The political uses of the Lisbon ‘script’ in European higher education policy. *Journal of European Public Policy*, 18(4), 584–606. <https://doi.org/10.1080/13501763.2011.560490>
- Cone, L., & Brøgger, K. (2020). Soft privatisation: Mapping an emerging field of European education governance. *Globalisation, Societies and Education*, 18(4), 374–390.
- Corbett, A. (2011). Ping Pong: Competing leadership for reform in EU higher education 1998–2006. *European Journal of Education*, 46(1), 36–53.
- Corbett, A. (2012). Principles, problems, politics... What does the historical record of EU cooperation in higher education tell the EHEA generation? In *European higher education at the crossroads* (pp. 39–57). Springer, Dordrecht.
- Curaj, A., Scott, P., Vlasceanu, L., & Wilson, L. (Eds.). (2012). *European higher education at the crossroads: between the Bologna process and national reforms*. Springer Science & Business Media.
- Curaj, A., Matei, L., Pricopie, R., Salmi, J., & Scott, P. (2015). *The European higher education area: Between critical reflections and future policies* (p. 898). Springer Nature.
- Curaj, A., Deca, L., & Pricopie, R. (2018). *European higher education area: The impact of past and future policies* (p. 721). Springer Nature.
- Dakowska, D. (2019). Higher education policy in the European Union. In W. R. Thompson (Ed.), *Oxford research encyclopedia of politics*, Oxford University Press.
- Dakowska, D. (2020). Let us make experts in our image: The European Commission and higher education policies. *Politix*, 2, 137–164.
- Dale, R. (2007). Changing meanings of “the Europe of knowledge” and “modernizing the university”, from Bologna to the “New Lisbon.” *European Education*, 39(4), 27–42.
- E4, (2011). Mapping the ESG.
- European Commission. (2006). *The history of European cooperation in education and training: Europe in the making—an example*. European Communities.
- Commission, E. (2017). *Strengthening European identity through education and culture—The European commission’s contribution to the leaders’ meeting in Gothenburg* (p. 11). European Commission.

- Commission, E. (2022). *Communication on a European strategy for universities*. European Commission.
- Enders, J., & Westerheijden, D. F. (2014). Quality assurance in the European policy arena. *Policy and Society*, 33(3), 167–176.
- ENQA. (2021). *Annual Report 2020*. ENQA.
- ESG. (2015). *Standards and Guidelines for Quality Assurance in the European Higher Education Area*, Brussels, Belgium.
- ETER. (2022). Website page. <https://www.eter-project.com/project-description/partner-institutions/>
- EQAR. (2020). *Use and interpretation of the ESG for the European Register of quality assurance agencies*. EQAR.
- EQAR. (2021). *Self-evaluation report*. EQAR.
- European Commission (2022). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on a European strategy for universities, SWD2022, 6 final. <https://education.ec.europa.eu/document/commission-communication-on-a-european-strategy-for-universities> [23.11.22]
- Fourcade, M. (2016). Ordinalization: Lewis A. coser Memorial award for Theoretical agenda setting 2014. *Sociological Theory*, 34(3), 175–195.
- Fumasoli, T., Stensaker, B., & Vukasovic, M. (2018). Tackling the multi-actor and multi-level complexity of European governance of knowledge: Transnational actors in focus. *European Educational Research Journal*, 17(3), 325–334.
- Fontdevila, C. (2023). The politics of good enough data. Developments, dilemmas and deadlocks in the production of global learning metrics. *International Journal of Educational Development*, 96. <https://doi.org/10.1016/j.ijedudev.2022.102684>
- Grek, S. (2013). Expert moves: International comparative testing and the rise of expertocracy. *Journal of Education Policy*. <https://doi.org/10.1080/02680939.2012.758825>
- Grek, S. (2014). OECD as a site of co-production: European education governance and the new politics of policy mobilization. *Critical Policy Studies*, <https://doi.org/10.1080/19460171.2013.862503>
- Grek, S. (2016). Knowledge actors and the construction of new governing panoramas: The case of the European Commission's DG education and culture. *Educação & Sociedade*, 37(136), 707–726.
- Grek, S. (2020). Prophets, saviours and saints: Symbolic governance and the rise of a transnational metrological field. *International Review of Education*. <https://doi.org/10.1007/s11159-020-09844-z>

- Grek, S. (2022). The education Sustainable Development Goal and the generative power of failing metrics. *Policy and Society*, 41(4), 445–457. <https://doi.org/10.1093/polsoc/puac020>
- Gornitzka, Å., & Stensaker, B. (2014). The dynamics of European regulatory regimes in higher education—Challenged prerogatives and evolutionary change. *Policy and Society*, 33(3), 177–188.
- Hartmann, E. (2017). Quality assurance and the shift towards private governance in higher education: Europeanisation through the back door? *Globalisation, Societies and Education*, 16(4), 309–324.
- Huisman, J., Adelman, C., Hsieh, C. C., Shams, F., & Wilkins, S. (2012). Europe's Bologna process and its impact on global higher education. *The SAGE handbook of international higher education* (pp. 81–100).
- Jarvis, D. S. (2014). Regulating higher education: Quality assurance and neoliberal managerialism in higher education—A critical introduction. *Policy and Society*, 33(3), 155–166.
- Krejsler, J. B., Olsson, U., & Petersson, K. (2012). Governing Europe by comparison, peer pressure & self-interest: On the Bologna stocktaking process as operator of national education policy. *Bulletin of Institute of Technology and Vocational Education*, 9, 35–47.
- Lawn, M. (2011). Standardizing the European education policy space. *European Educational Research Journal*, 10(2), 259–272.
- Lawn, M., & Grek, S. (2012). *Europeanizing education: Governing a new policy space*. Symposium.
- Magalhães, A., Veiga, A., Sousa, S., & Ribeiro, F. (2012). How is European governance configuring the EHEA? *European Journal of Higher Education*, 2(2–3), 160–173.
- Montoya, S., & Crouch, L. (2019a). The learning assessment market: pointers for countries-part 1, UNESCO World Education Blog. <https://world-education-blog.org/2019/04/26/the-learning-assessment-market-pointers-for-countries-part-1/> [accessed 19.11.23]
- Montoya, S., & Crouch, L. (2019b). The learning assessment market: pointers for countries-part 2, UNESCO World Education Blog. <https://world-education-blog.org/2019/05/20/the-learning-assessment-market-pointers-for-countries-part-2/> [accessed 19.11.23]
- OECD. (2022). *Labour Market Relevance and Outcomes of Higher Education*. OECD.
- Ollikainen, A. (1999). *The single market for education and national education policy*. Research Unit for Sociology of Education.
- Power, M. (2003). Evaluating the audit explosion, *Law and Policy*, 23(3), 185–202.
- Ozga, J., Dahler-Larsen, P., Segerholm, C., et al. (Eds.). (2011). *Fabricating quality in education: Data and governance in Europe*. Routledge.

- Robertson, S. L. (2008). Embracing the global: Crisis and the creation of a new semiotic order to secure Europe's knowledge-based economy. In B. Jessop, N. Fairclough, R. Wodak (Eds.), *Education and the knowledge-based economy in Europe* (pp. 89–108). Sense.
- Sarakinioti, A., & Philippou, S. (2020). European discourse on higher education quality assurance and accreditation: Recontextualizations in Greece and Cyprus at times of 'Crisis.' *European Education*, 52(2), 132–145.
- Scott, C. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.
- Shaw, J. (1999). From the margins to the centre: Education and training law and policy. In P. Craig & G. de Búrca (Eds.), *The evolution of the EU law* (pp. 555–595). Oxford University Press.
- Sorensen, T. B., & Robertson, S. L. (2020). Ordinalization and the OECD's governance of teachers. *Comparative Education Review*, 64(1), 21–45. <https://doi.org/10.1086/706758>
- Schriewer, J. (2009). Rationalized myths" in European Higher education: The construction and diffusion of the Bologna Model. *European Education*, 41(2), 31–51.
- Stensaker, B., Harvey, L., Huisman, J., Langfeldt, L., & Westerheijden, D. F. (2010). The impact of the European standards and guidelines in agency evaluations. *European Journal of Education*, 45(4), 577–587.
- Veiga, A. (2019). Unthinking the European Higher Education Area—differentiated integration and Bologna's different configurations. *Innovation: The European Journal of Social Science Research*, 32(4), 495–512.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Quantification and Utopia: New Forms of Expert Knowledge Production in Global Education Governance

1 INTRODUCTION

In the first chapter of this book, I discussed that the starting point for the METRO project was an investigation of the emergent collaborations of International Organisations in the field of the production of metrics for transnational governance. The research project focused on an analysis of four case studies, on the measurement of education (on its global and European dimensions), on global poverty and on the rise of statistical capacity development, especially of statistical offices in the Global South. I described how, from the very start of the fieldwork, the notions of interplay and interdependence were instrumental to the analysis. In particular, what became quickly apparent was that quantification was not only made through the collaborations of large international organisations, but also that numbers had become the connective tissue of a large and ever-expanding governing architecture, what I and my colleagues have elsewhere described in detail as an ‘epistemic infrastructure’ of transnational governance (Bandola-Gill et al., 2022; Grek, 2022; Tichenor et al., 2022).

As a result, especially in the policy arena of education, METRO was dominated by the study of quantification in two major monitoring and policy initiatives: these were the Sustainable Development Goals (SDGs) and the making—for the first time ever—of the European Education Area (EEA). The SDGs, despite the slow progress towards their achievement,

have been transformational in that they have truly created a global education policy space, where all participating countries, both from the North and the South, are monitored against their performance in a diverse range of indicators, from learning outcomes in literacy and numeracy, to gender equity and citizenship. In addition, for the first time ever, there has been a substantial shift in the geopolitics of the influence of large IOs in the field of education. Their collaborations and synergies have led to a much more fluid space of interaction, where older notions of a certain ‘territorialisation’ of zones of influence (OECD in the North, World Bank and UNESCO in the South) do not appear to hold ground any longer. Finally, as Chapter 6 outlined, the construction of a European Education Area is the first open proclamation of a unified strategy for the making of European education as a single policy arena. Despite the rule of subsidiarity that adjudicates education as a national matter (and thus not in the jurisdiction of the EU), the dominance of datafied governance across education stages and institutions has now allowed the European Commission to overcome political sensitivities, and support financially and strategically the construction of education as one of the key policy fields in the EU; numbers, of course, and all the calculative rationalities they created, have been the main motor powering this political and symbolic shift.

Although analysing quantification as the rise of an epistemic infrastructure—where the materiality of data is entangled with actors and networks (Bandola-Gill et al., 2022; Grek, 2022)—was a useful way to understand the paradigmatic changes that are taking place in the relationship of knowledge-making with governance, *The New Production of Expert Knowledge* is a fine-grained analysis of the constitutive qualities of expertise in the twenty-first century. In order to allow for such an analysis, the notion of the Mode 1 to Mode 2 knowledge transformation was key, as it was against its constituent parts that the findings of METRO were compared and contrasted. Although the idea of Mode 2 has been characterised as ahistorical, normative and ultimately part and parcel of the ‘governance turn’ at the end of the twentieth century, Mode 2 represented a key moment in re-thinking the relationship of knowledge production with governing, and thus became a popular idea that made traction not only in its own STS field but further afield, and particularly in the sociology of expertise. It is thanks to its usefulness as a thinking aid that it has taken such a central role in this book, breaking down the different constituent elements of the shifts in knowledge production and

guiding the writing of each chapter. In what follows, I will return to these elements to discuss them in turn, before moving on to theorising how the new production of expert knowledge has become a key ingredient of educational utopia-making.

2 FROM MODE 2 TO THE PRODUCTION OF NEW EXPERT KNOWLEDGE IN THE TWENTY-FIRST CENTURY

Before moving to the analysis of new developments in expert knowledge production, I would like to return to the outline of the ways Mode 2 worked as a productive springboard for thinking and analysing the making of expertise in the transnational governance of education:

<i>Mode 1 Knowledge</i>	<i>Mode 2 Knowledge</i>	<i>Expert Knowledge</i>
University context Disciplinarity	Context of application Transdisciplinarity	Global/universal level Post-disciplinarity/ Mono-disciplinarity (Economisation)
Homogeneity	Heterogeneity and Organisational Diversity	Brokerage/consensus/mediation
Autonomy	Social accountability and user reflexivity	Datafied accountability and expert reflexivity
Peer-review quality control	Extended quality control	The market of measures

Although the table format may imply an evolutionary or even comparative element to the relationship of Mode 2 and the making of expert knowledge, I do not wish to claim that expert knowledge qualities, as identified in the third column above, represent the further evolution of knowledge *beyond* Mode 2; the table is used merely as a heuristic schema which, although acknowledging that other contexts and spaces may display different characteristics in their knowledge/ policy relationship (Mode 2 or even Mode 1), at least in the field of the European and global governance of education, the story of expert knowledge production is substantially different than the one narrated by Mode 2. We will now look at these qualities in turn, in an attempt to make better sense of them.

2.1 *Global Data for Universal Values*

Despite the critique, the concept of Mode 2 science entailed an important point in theorisations of the relationship between knowledge and governance, which could be reflected in the debates over the focus on knowledge production in the *context of application*. The idea of producing knowledge across dispersed stakeholder groups and across multiple and diffuse institutional boundaries gained considerable traction as an antecedent of the idea of ‘co-production’ of science and policy. This idea was taken up by Sheila Jasanoff (2004) and is still relevant and increasingly popular in the production of research, and social science research in particular:

Increasingly the realities of human experience emerge as the joint achievements of scientific, technical and social enterprise: science and society are, in a word, co-produced, each underwriting the other’s existence. (Jasanoff, 2004, p. 17)

The idea of producing knowledge at the context of application entails a construction of knowledge that renders the boundaries between research users and producers insignificant (Wyborn et al., 2019). With co-production, this blurriness goes further, with the research process being designed collectively among producers and users (Bandola-Gill et al., 2023). Such knowledge production requires careful navigation between different values, objectives and epistemic frameworks in order to both assure the pluralism of voices involved in the process of knowledge co-production (Lövbrand, 2011) but also its ‘usability’ (McNie, 2007).

After the ‘participatory turn’ of major international bodies and institutions, the study of the production of expert knowledge for the transnational governance of education, poverty and statistical development revealed a space of interdependence and collaboration that for the first time brought together not only technical experts, but also a much larger group of data producers, users, donors and national officials. These linkages and new entanglements encapsulate the enhanced role of a growing number of actors—from the UN agencies and member states, to philanthropic and civil society organisations and the academia, all participating in the production of global governance, not only at the decision-making level, but also in technical meetings and exchanges. These complex relationships happened through and around numbers, which both stabilised the connections, as well as mobilised and created

new constituencies and new interdependencies. Given the enhanced role of diverse actors in technical meetings and decision-making for determining the terms and conditions of ‘governing by numbers’, quantification emerged as a fruitful arena for collective puzzlement, socialisation and policy translation. Although in most cases it is the objectivity of numbers that is considered central for ‘governing at a distance’ (Cooper, 1998), governing the global education policy field intersected with two other elements; these are the notions of *symbolic space* and *belonging*. Space is crucial, because the transnational participatory turn, contrary to other sites of audit and accountability, necessitated meetings at specific physical and online places. This is important to keep in mind, since often the discussion around numbers, standards and performance management appears as relatively abstract and top-down, therefore, missing out on an understanding of the role of meetings that bring together a community of people. On the other hand, we saw that progressively—and even in politically and historically distinct, at times even hostile, organisations, such as the relationship between the OECD with the European Commission, or the World Bank with UNESCO—socialisation led to *belonging*; this concept was relevant to the analysis, as contestations were counterbalanced by a sense of universality and rapport, ‘mobilised by institutions in their struggles over acceptable political practice’ (Cooper, 1998, p. 16).

Still, how can one contemplate that the cold rationality of number-making could ever lead to such collective declarations of belonging? The book charted the significance of numerical inscriptions in the production of shared narratives and global values. Chapter 2 discussed how global and European monitoring exercises are not merely technical exercises but discursively address some of the most complex, interlinked and compounded global challenges that the world currently faces. Indeed, dystopian numbers became affective tools in the arsenal of persuasion devices for IOs’ experts. Therefore, the production of numerical 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, numerical narratives do not need to be precise; thus, they offer added legitimacy to numbers, while masking data gaps and technical inaccuracies. According to Roe, 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 recognisably so—but still persist, indeed thrive’ (1994, p. 51).

Above all, in the field of the global governance of education, contrary to the Mode 2 argumentation, contextualization (or the context of application) does not appear relevant. Instead, what we see is the articulation of global values into local ones. In the space of the global governance of education, extended as it has been with the participation of national and local actors and agendas, the local is erased, as it is translated into—or sidelined by—global, universal values that bind all actors, numbers and narratives together in a discursive mix. In other words, and perhaps counterintuitively, the global *has become* the local and vice versa: this is the new doxa of the universality of education problems and solutions that appears to guide most of the production of expert knowledge and education policy in the twenty-first century.

2.2 *Monodisciplinary Visions of a Complex Education World*

One of the key transformations of Mode 2 knowledge production was the idea that knowledge had to break free from the siloed disciplinary confines and be synthesised with other knowledge in order to be relevant and effective: this was the move towards transdisciplinary knowledge production, which gained particular traction in the evidence-based policy literature, as almost a pre-condition of solving the wicked super-crises beholding societies. As if transdisciplinarity were not hard to achieve in the first place, most knowledge producers (and funders) moved on to an emphasis of the benefits of *interdisciplinarity*, as the need not only to bring together but also to synthesise and combine disciplinary perspectives grew (though, admittedly this was not what Mode 2 proclaimed). According to the supporters of interdisciplinary knowledge production, global challenges require the combining of methods and insights from multiple academic disciplines in order to resolve the multifaceted and complex ‘wicked’ problems of the twenty-first century, such as inequalities or sustainability. Education, with its close relevance to a number of social processes, problems and opportunities (for example, its links to citizenship, democracy, sustainability, labour markets, inequalities, well-being, health, innovation and others) were to benefit most from such a synthesis of expertise in order to address complex social issues; as we have seen, however, this is not what happened.

Instead of trans/interdisciplinary knowledge, the production of new expert knowledge in the field of the global governance in education is characterised by *mono-disciplinarity*. This is the rise and dominance

of economics, as the only relevant disciplinary field that would bring such diversity of actors and their interests together. The supremacy of economic versus any other knowledge in the global governance of education is of course not new. It can easily be traced and explained by the key role that at least two major international organisations played in the formation of the field of transnational education and the education indicators and data that shaped it: these are the Organisation for Economic Cooperation and Development (OECD) and the World Bank, both well-known for the close links that they draw between education, economic growth and human capital development. Similarly, the European Commission, despite its support for education for the promotion of social cohesion and a ‘people’s Europe’ (Grek, 2008; Lawn & Grek, 2012), has not shied away from its emphasis and support for education as a key driver for the making of a prosperous and competitive ‘global Europe’.

To be clear, my argument does not refer to the economisation of education as the well-rehearsed analysis of neoliberalism and its effects on education over three decades and more; although this form of critique is still relevant, it has been developed eloquently in previous research (Barrett, 2011; Mundy et al., 2016; Tikly, 2015). What monodisciplinarity means in the context of the global governance of education, is the sole dependence of education as a policy field on economic epistemology and methodology as a way of mapping, knowing and planning education for the future. Although arguably this dependency cannot be decoupled from the historical roots of the construction of a commensurate global education policy field by international organisations that prioritised economic perspectives (and specifically neoliberal economic values) (for a comprehensive analysis, read Elfert & Ydesen, 2023), it is important to pay attention to how economics shapes the production of expertise in education and what the effects are for knowing and governing the field.

As discussed in Chapter 3, although interviews with education experts revealed some differences among them, they all shared one common characteristic: that is, their disciplinary background was in economics. In the early decades of the production of education indicators (1970s and 1980s), it was education economists that pushed for the idea of building comparative education datasets, in order for major economies to compare and compete in terms of education performance for economic prosperity. Later, at the turn of the century and into the 2000s, as we saw in

Chapter 3, it was again education economists, such as Hanushek (2000) and Glewwe (2002) that led to the paradigmatic shift from the measurement of inputs to outputs in education (Grek, 2022). As METRO's fieldwork revealed, many of their disciples continue to dominate the field; although working in the field of education, when asked about their studies, they all responded that they had economics degrees.

Although education economics is a needed and long-standing way of doing comparative analysis in education, it is the singularity and prevalence of economics as a vision, a way of thinking and analysis that is of interest here: following its epistemology, education processes and institutions are framed as economic entities with unlimited possibilities for growth and improvement (Miller & Power, 2013). A focus on economics as a way of structuring and comparing education data 'implies a concern with the idea of efficiency (Kurunmäki et al., 2016, p. 396), as well as the aim to create and expand the education 'market' (as we saw in Chapter 6), and an emphasis on competition and performance (Caliskan & Callon, 2009). Last but not least, in education but also a multiplicity of other areas of political and social life, the prevalence of economics and the economisation of knowledge production that followed it, has led to the financialisation of education actors, processes and institutions, calculated as assets in a capital investment market that is hoped will create returns (Chiapello, 2015; Muniesa et al., 2017); post-COVID, such processes of assetization have intensified following the increasing digitalisation of education and its services.

In a world of increasing and compounding challenges, how come the global governance of education has become so dependent on the mono-disciplinarity of economics? It is in the role of quantification that the answer has to be found, in what looks increasingly like a chicken and egg question. Mennicken and Espeland are beautifully eloquent in their description of the relationship of quantification with economics; their analysis could easily work as an accurate description of the education condition for more than half a century now:

Quantification and commensuration are key conditions for economic calculation and action. Quantification makes individual and organizational performance visible, trackable, and comparable, thereby allowing for organizing in accordance with principles of efficiency. (2019, p. 240)

2.3 *Perfect Brokers of Imperfect Numbers*

Mode 2 knowledge proponents suggested that, as social problems become increasingly ‘wicked’, no one source of expertise is sufficient to solve policy issues (Baekkeskov, 2016). Consequently, as highlighted by Nowotny et al. (2003, p. 155), new challenges require ‘socially distributed expertise’—one which is decentralised and blends multiple sources of knowledge and actors sourcing evidence. Such plurality of expertise was not seen as limited to elite knowledge producers, but would also include localised ‘lay’ experts and experts outside of academia, where the traditional scientific knowledge was produced.

At the same time, the key tension in identifying expertise in high-risk settings cannot be reduced to simply adding new forms of expertise to the equation, but rather requires the emergence of new, different forms of expertise altogether (Eyal, 2019). Such transformations to the nature and role of expertise were also a result of what was seen as its high-context relevance and thus its inherently ‘local’ nature (Wyborn et al., 2019)—a characteristic that was also understood to be contradicting the rise of increasingly global challenges.

Although the METRO interviews revealed a diverse field of actors participating in the production of monitoring agendas, IOs’ experts have continued to occupy a central role. Their expertise consisted primarily in the evaluation and harmonisation of datasets, as the latter were produced by national and international assessments, and more crucially in their ability to use their epistemic capital, as well as their socialisation skills in order to broker knowledge between actors and fields and persuade participants about the benefits of their involvement. After the participatory turn that the SDGs brought, and following similar developments at the European education policy field where measurement agendas have to be politically acceptable to all member states involved, expert work has evolved to include more than simply the production of robust data. Instead, the principles of democratisation and technocracy are considered indivisible and thus leading experts to apply their mediation skills in order to first, secure country ‘buy-in’ into the monitoring frameworks; second, navigate local politics and requirements, especially in cases of countries of the global South where needs for the collection of global, comparable data do not match local budgets and needs; and finally, to succumb some of their technical robustness to the politics of producing ‘good enough’

data (Fontdevila, 2023), so as to allow the process to continue and a level of minimum consensus to be found.

Such brokering and mediating work depended on experts working to maintain a tight balance between retaining their epistemic superiority and scientific credentials, while combining these virtues with the crucial political calculations of how to secure (and maintain) participation and buy-in, create consensus and make sure that the numbers' work continues apace: that is, despite the imperfections of 'bad' quality data, experts have the—often impossible—task of producing indicators that are acceptable to governments, fit their existing local data structures, while also being comparable globally and reaching a minimum level of technical quality. Dealing with the implications of having to balance out the technical *and* the political challenges of doing expert work, experts' brokering practices were aided by the use of 'imperfect' numbers (for a fuller discussion of the distinction between 'ambiguous', 'placeholder' and 'provisional' numbers see Chapter 5 in Bandola-Gill et al., 2022) which, as we saw in Chapter 4, counterintuitively transformed IO actors into more trusted experts than those who depended solely on their epistemic authority to maintain their influence and trustworthiness in the field. Although it is quantification that brings all actors, narratives and numerical inscriptions together, expert work has transformed to adapt to a much broader conceptualisation of what epistemic authority entails; that is, moving beyond the objectivity and scientific robustness of numerical work, quantification, through the experts' work that carry it, appears to be folding both the science and the politics of numbers in its processes, thus expanding and re-inventing what 'trust in numbers' in the twenty-first century entails.

2.4 *Reflexive Experts of Datafied Systems*

In tandem with the making of knowledge within the context of application, Mode 2 suggested that contemporary knowledge production cannot be autonomous any longer (knowledge for knowledge's sake) but has to be socially accountable and reflexive: in other words, knowledge producers need to be accountable to the communities that they belong, i.e. produce knowledge that is understandable and justifiable by its users. Further, it was understood that such socially accountable knowledge production would also lead to increasing levels of user reflexivity, in relation to the knowledge produced.

Nonetheless, what the last three decades in the field of the global governance of education have shown is not the emergence of socially accountable expert knowledge production, but rather a much more technical, performance-based accountability that saw the rise of an assemblage of formal and informal procedures, various techniques, assessments, tools and normative discourse, aiming at making education systems accountable (by making them comparable) for improving performance. This form of performance-based accountability is dependent on the prevalence of datafication, that is, on quantified data that originates from local and national testing, other forms of evaluation and comparison and is eventually translated into performance indicators and global measurements of learning data. Datafied accountability can thus connect and interlock several scales, from local and sub-local levels, to the national and the global and is the key ingredient of the rise of ‘expertocracy’ in education (Grek, 2013). From the 1980s/9s New Public Management and the emergence of high stakes testing, all the way to the rise of international comparative assessments and the datafication of education, performance-based accountability is the outcome of historic, contemporary and social constructions that result in the selection, bricolage and translation of datafication in education policy globally (for a more extensive discussion of accountability and datafication, please see Grek et al., 2021).

The rise of datafied accountability has become the day-to-day reality of all education systems around the world, with severe repercussions on schools, students, parents and the teaching profession. It is a story that education researchers have told and, as the METRO findings showed, one that has increasing impact on experts’ reflexivity, as they observe the limitations, challenges and often the countereffects of almost half a century of education datafication. Chapter 5 discussed how surprised the research team was to find such a heightened sense of reflexivity among experts, who openly discussed not only the difficulties of their task, but also the political nature of data production. During the 100 + IO expert interviews that we took, we found increasing numbers of experts who were revealing about the challenges of bringing together technocratic and participatory modes of decision-making. Experts discussed the difficulties of finding appropriate data and datasets to match the selected indicators, given that the latter were not decided on the basis of what was technically possible but were merely political aspirations and declared policy goals; many of them shared their frustrations of the ‘endless’ consultation processes that took over their technical meetings. Lastly, as

I discussed, experts in the field, and particularly those that aimed to capitalise from their long-standing relationships with countries in the global South, rather than their technical credentials, used reflexivity instrumentally, proclaiming the benefits of pushing the agenda on, even if they had to do so on the basis of imperfect or approximate data. Indicators, as they suggested, were never meant to be accurate; they are to be used as ‘barometers’ of countries’ progress rather than precision tools (Montoya, 2017).

Thus, instead of social accountability and user reflexivity, expert knowledge production in the twenty-first century has for a long time now followed a path of performance management and datafication; the repercussions of these trends are reflected in the reactions against them, leading to what I have discussed as the increasing democratisation and the participatory turn in global governance. Through such processes of widening the field of actors, of democratising the agenda and of using numerical narratives and affective scripts, expert knowledge producers have become a lot more reflexive of their work and its challenges. Using epistemic reflexivity to talk about the unintended consequences of quantification, or even applying reflexive practices instrumentally in order to enhance their persuasive power, experts appear more and more to be producing knowledge because they *care* for the communities and the issues they work on. They may be technocrats, with specific educational paths and career trajectories, but they can also be ‘prophets, saviours and saints’: that is, they measure the present and forecast the future with the aim to save lives, despite the challenges and out of a sense of morality and altruism towards those mostly in need (as I have discussed elsewhere, see Grek, 2020). Therefore, the opening out of the field of the transnational governance of education revealed an expert knowledge production that is highly reflexive, dialogic, self-critical and open to contestation.

2.5 *The Market of Measures*

Mode 2 suggested that the quality control of new knowledge production was not only limited to the traditional peer review mechanisms that scientific knowledge has always been submitted to, but also that the quality control was more extended and dispersed, including its users and the wider public. Given the move to more transdisciplinary and contextualised knowledge production and application, such a shift to how the quality of

knowledge was assured was seen as the direct consequence of the wider transformations taking place in society.

Expert knowledge production in education has also seen such a broadening out of the quality controls that establish its worth. Although it experienced a long period of dominance and success of certain measures (see OECD's PISA for example, which ruled media headlines and education ministers' desks for at least a decade), what we see as a direct consequence of the production of SDG4, is a competition of measures. Most IOs, although working collaboratively, are also eager to promote their own institutional brand and thus operate simultaneously at two levels: on the one hand, they appear as open to alliances and working collaborations towards 'the goals', while on the other they are also conscious of needing to maintain their independence and unique contribution to the field. These double roles create contestations and lack of trust and may derail negotiations in what is already a very fragile governing field. Within Europe, the market of measures is also growing: here, I examined the production of quality assurance in higher education in Europe and showed how a project that began developing around the Bologna process twenty years ago, has now grown into a fully blown quality assurance industry, with a growing number of actors and datasets competing for attention and funding. As some of the key figures in the global governance of education discuss (see Chapter 6), this is a market of measurement, with inefficiencies and paradoxes, but also sellers, buyers and sale pitching events for determining which 'product' best fits the work of policymakers. Given the influence of private capital in education and—especially post-COVID—the proliferation of major education consultancies and platforms, the market of measures and data producers is only bound to inflate and become more competitive. However, Chapter 6 does not ask whether there is a market of education data providers or not; we have known for some time that it exists and that it is thriving. Instead, the focus of the argument is that, instead of collaboratively working towards producing the most robust measures for the calculation of global learning data, selecting measures appears more like a 'pick-n-mix'; governments and IOs select the measures best suited to the budget, their policy priorities and above all, the way the measures portray their performance.

To conclude this section, the transformation from Mode 1 to Mode 2 knowledge production promised a much more open and horizontal field, where knowledge was to be applied, contextualised, transdisciplinary

and led by user demand. Consequently, an examination of the qualities of expert knowledge has become increasingly complex, as the criteria of assessment are dispersed and often contradictory depending on the stakeholder. The moves towards Mode 2 knowledge production have led not only to the democratisation of knowledge by increasing the plurality of actors involved in the process of knowledge production (Nowotny et al., 2001), but rather also to the commodification of knowledge and its assessment purely in utilitarian sense (Ozga et al., 2011). Yet, despite the tensions between the production of authoritative, ‘usable’ expert knowledge and the critical, up-stream engagement of stakeholders with it (Lövbrand, 2011), quantification persists as the only viable means by which to plan and prepare for a better world of free and equal education for all. Despite the fairly bounded, national responses to the major education crisis that the recent global pandemic brought to the fore, it is clear that COVID-19 has acted as an accelerator for the re-making of datafied and digitised education governance in the twenty-first century. Building on discourses around the ‘devastating impact’, ‘learning losses’ and yet another ‘lost generation’ that COVID-19 may have brought (Brookings, 2023), global education experts try to tame the current condition of radical uncertainty by inscribing the future into calculable horizons. The next and final section of this book discusses the role of quantification in the promise and the crafting of utopian futures of education that have never been.

3 EDUCATION, QUANTIFICATION AND UTOPIA

Utopias are not new in education. From Plato’s *Republic* to the critical pedagogy of Paulo Freire, the dream of creating an alternative society, free from oppression and inequality and guided by critical pedagogy, has been at the centre of educational thought and action for a very long time. In this concluding section of the book, I discuss how quantification, from a technocratic and a-political mode of informing policy with evidence, has slowly been transforming into a mode of utopian thinking—a way of seeing, constructing and performing the ‘desired possible worlds’ of the future (Levitas, 1990). Here, I do not examine utopian thinking as the practice of dreaming education unicorns and sunlit uplands, which is a frequent criticism of the term. Instead, I follow Jameson’s ‘Politics of Utopia’ (2004) and Levitas’ ‘Utopia as Method’ (2013) to reflect on the ways that quantification has not only adopted elements of forecasting

and planning ideal future worlds, but also that the processes of number-making have acquired the social and political function of co-constructing alternative and utopian education panoramas.

In particular, this book has built on theorisations of numbers as performative, i.e. constructing rather than simply measuring political phenomena (Kingsbury et al., 2012; Mehrpouya & Samiolo, 2016; Porter, 1995) in order to show how, quantification is increasingly integral to the making of contemporary utopian education thought. By using the frame of ‘utopia-making’, I refer to the predominance of the co-construction of ambitious political and education imaginaries via numbers. The analysis of the making of the SDGs, as well as the production of a single and unified European education area, reveals numbers as central to the making of utopian visions of interconnected education policy futures, rather than merely representing sets of isolated targets and policy recommendations, as previous global tools like PISA or the MDGs involved.

In particular, this book has shown how older arguments about the de-politicisation of numerical work have now been replaced by its *re-politicisation*, as the politics of numbers are not hidden any longer, but legitimised on the basis of their transformation into the new spaces for democratisation of decision-making. Elsewhere I have charted the ways that the rise of an epistemic infrastructure in global public policy has led to the paradigmatic shift of the recalibration of measurement and governing as co-constructed: measurement and the production of expertise are advocated and utilised as key spaces for achieving political consensus and for shaping public policy directions (Bandola-Gill et al., 2022; Grek, 2022). In a post-truth world, flooded with data and mistrusted numbers, expertise had to transform and adapt to these new political and social realities; these developments were strengthened further through social movements such as ‘Black Lives Matter’ (Strickland, 2022), #MeToo (Hillstrom, 2019), the rise of decolonial discourses (Bhambra, 2014) and the threat of climate change. Thus, as the pages of this book have revealed, expert knowledge production has assumed functions that would have been previously unthinkable. Instead of merely informing policy, expertise has become the platform for envisioning new ways of doing governing: interpretative flexibility, openness, (re)politicisation, reflexivity and democratisation are key discourses and proclaimed aims for the new governing *and* expert knowledge paradigms of the twenty-first century. Thirty years on from *The New Production of Knowledge* (Gibbons et al.,

2010), quantification is not merely a tool in the arsenal of policy instrumentation and change; instead, quantification has been institutionalised as being at the very core of governance itself.

Although these phenomena and their management are of high significance in contemporary governance practices, here I analyse specifically the role of numbers as the key building block in the making of utopian education futures. Indeed, as we have seen in the pages of this book, measurement and the making of expert knowledge more broadly, have moved beyond achieving mere prediction and ‘readiness’. Rather than offering the reading of a crystal ball, narrating the future involves establishing a discursive agenda of the education values of the present, and the ideas and ambitions of how education futures will be shaped. In bringing together these futuristic ambitions and goals (‘ensuring inclusive and equitable education for all’ for example—the education SDG), narrating the future represents a governing manifesto of contemporary considerations, uncertainties and potentialities. By quantifying these education futures, actors in European and global education governance establish a common utopia of political goals, the progress of which can be carefully measured and supported by policy reforms.

However, how can the post-COVID, crisis-ridden education policy arena have fostered such ambitious education plans? The book, through its analysis of numerical narratives, networks of actors, meetings, data harmonisation processes and aspirational policy declarations, showed how quantification serves as the springboard for the ‘utopian leap’: that is, the quest for numbers counteracts the paralysis of a dystopian reality and fills the gap between the dreary present and the ideal arrangements of a desired future. Thus, the ‘promise and dream’ of quantified future-making is grounded on a central condition: utopia does not only offer the imaginary of an ideal world, but it is also—and crucially—inherently *procedural* (cf. Thaler, 2019): in other words, through the establishment of the processes of indicator making, data harmonising, actor meeting, report writing and many others, the work of producing the future is being done. This is what Ruth Levitas (2013) coined as ‘utopia as a method’—one that is not only about imagining better worlds, but also a mode of action. In addition, through the imagining of alternative realities, utopias necessarily encourage a reflection on the current state of the world. Indeed, recent scholarship on utopianism offers a more nuanced understanding of the relationship between utopias and reality. The most important work

aimed at merging the two, rather than contradicting them, is Erik Olin Wright's 'Real Utopias Project'. As argued by the author,

What we need, then, is 'real utopias': utopian ideals that are grounded in the real potentials of humanity, utopian destinations that have accessible waystations, utopian designs of institutions that can inform our practical tasks of navigating a world of imperfect conditions for social change. (Wright, 2010, p. 6)

As I argue in this book, the new expert knowledge production for global governance, re-imagined and re-organised as a common space for technical *and* democratic accountability, represents an example of 'utopia as method': by painstakingly drawing a multiplicity of indicators and actors together, they outline these 'waystations' for all participant countries as the only realisable—and available—path to a better future. Ruth Levitas suggests that the utopian vision may foster conditions for thought, debate and experimentation (1990): as I have shown, meetings and data collection practices ostensibly create spaces for exchange, no matter how unequal and asymmetrical they may be. Further, they also appear as helping facilitate criticism of the current reality—in the context of global governance, such criticism does not only relate to the state of the world per se, but also what has been seen as the continuous colonial project of the global North determining the future of the global South (see Boldero & Francis, 2002). As we have seen, the demands for democratisation and decoloniality have destabilised the older balance of power among IOs and created new opportunities for previously weaker IOs to gain new influence (as the example of the UNESCO Institute of Statistics has shown—see Chapter 5).

However, as the analysis showed, for quantification to represent a new mode of political imagination, utopias of imagined future education worlds have to be coupled with dystopian thinking. Dystopias are inherently grounded in a 'cautionary pedagogy'—a warning about the state of the world and its future. As such, dystopias play a specific role in galvanising actors involved in dealing with a crisis (Thaler, 2022). The goal of a dystopia is to imbue utopian visions' 'wishful thinking' into pragmatic and—in the case of quantification—technocratic and process-based modes of inquiry, thus rendering utopian visions as problem—rather than idea-oriented (cf. Gümüşay & Reinecke, 2021). Jameson could not have phrased this better: utopia-making represents 'model railroads of

the mind', the continuous 'bricolating and cobbling together things of all kinds' (Jameson, 2004). This is utopian thinking as 'miniaturization': that is, 'replicating ... things in handicraft dimensions that you can put together by yourself and test ... or change and rebuild in a never-ending variation fed by new ideas and information' (Jameson, 2004, p. 40).

Thus, it is the material and procedural character of utopian thinking that aligns it so closely with the analysis of expert knowledge production for governance. As I have shown in detail in the previous chapters, expertise, with its novel qualities of universalism, mono-disciplinarity, brokerage, reflexivity and marketisation, created a more fluid space that assembled a broad church of actors, ideas, methods and interests. The intertwining of these infrastructural elements came together, time and again, in smaller or bigger ways, to create perfect versions of an evidence-focused, goal-oriented, utopian future. By 'coming together' I do not intend to say that these processes were harmonious, frictionless or equal in any way. The analysis of the production of European and global learning data is a space of struggle and contention, where numbers facilitate debate and cause discord, but also that, at critical junctures of the process, worked to achieve at least a minimum consensus over goals and priorities, so as to protect the place of education as a key policy arena in the field of the global governance of sustainability and economic growth.

Perhaps the most notable affordance of quantification to create spaces of educational consensus has been the quality of numbers to be malleable and moving, rather than fixed, entities. Rather than valuable for their objectivity, collecting data for education indicators became a process of finding 'good enough' data solutions for the short term: these could be ambiguous numbers, or what interviewees called 'placeholder' or 'provisional' numbers—and their value rested primarily on their 'strategic ambiguity' (Sillince et al., 2012). Such ambiguity of the contested indicators enabled them to act as boundary objects, almost 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 among them, but facilitating the continuation—and even bolstering—of number-making. This process of widening participation in the decision-making around indicators and quantitative targets is a focal point not only as a matter of achieving equity, but also—and perhaps primarily—as a route for enhancing political buy-in into the infrastructure of measurement. Therefore, the previously technocratic process of developing and validating indicators has transformed into a

forum for the construction of socio-technical imaginaries of a common, utopian future, carefully balanced between idealistic orientations, but also concerned with realism-driven expectations of feasibility. It is this interplay between idealism and pragmatism that quantification has achieved; trust in numbers bridges the promise of accountability and scientific authority with the political demand to create the—so-called at least—‘bottom-up’, collective and ambitious futures for the planet.

To conclude, despite their failures, European and global monitoring agendas and their expert producers have not lost their relevance, as they continue to dominate the global political debate around the need to ameliorate the chronic neglect and exploitation of both the environment and of vulnerable populations around the world. The exceptional global temperatures, rainfall and draughts of recent years are only signals that the global pandemic may not have been the worse humanity has experienced in the first half of the twenty-first century. This book discussed the ways the new production of expert knowledge represents social and political endeavours to rationalise and ‘technicize’ the process of offering education for all. It also discussed the ways quantification as utopia has overshadowed and monopolised any other spaces and modes of political imagination, as it proclaimed to offer technical and measurable ‘waystations’ in imagining and planning for education futures. This is because, as Miller contends,

the future does not exist in the present but anticipation does... To use the future is strictly speaking, not possible, since the future does not exist as an object or tool to be used. The future as anticipation, however, is continuously instrumentalised. (2018, p. 59)

Thus, quantification in global governance captured the imagination of a wide set of actors, since it purposefully allowed multiple ‘entry points’ in its world: although experts continued to emphasise the use of technocratic and management principles to create an objectified and measurable field (and hence still appear authoritative, accountable and in line with scientific approaches to policymaking), quantification is equally now being proclaimed as the space for bottom-up, grass-roots and transformative education politics. Despite power differentials, as well as tensions and disagreements, a common global education policy field has been created, making quantification the common policy language in the process. Ultimately, as we have seen, the production of quantified utopian futures,

as the only option to avoid catastrophic dystopias, may have little to do with the future itself: rather, it offers productive tools to make sense of and tame an increasingly ungovernable, crisis-prone and fast-moving educational present.

REFERENCES

- Baekkeskov, E. (2016). Same threat, different responses: Experts steering politicians and stakeholders in 2009 H1N1 vaccination policy-making. *Public Administration*. <https://doi.org/10.1111/padm.12244>
- Bandola-Gill, J., Arthur, M., & Leng, R. I. (2023). What is co-production? Conceptualising and understanding co-production of knowledge and policy across different theoretical perspectives, *Evidence & Policy*, 19(2), 275–298. Retrieved July 17, 2023, from <https://doi.org/10.1332/174426421X16420955772641>
- Bandola-Gill, Ju., Grek, S., & Tichenor, M. (2022) *Governing the sustainable development goals: Quantification in global public policy*. Springer International
- Barrett, A. M. (2011). A Millennial Learning Goal for education post-2015: A question of outcomes or processes. *Comparative Education*, 47(1), 119–133.
- Bhambra, Gurminder K. (2014) *Postcolonial and decolonial reconstructions in connected sociologies*, Bloomsbury Academic
- Boldero, J., & Francis, J. (2002). Goals, standards, and the self: Reference values serving different functions. *Personality and Social Psychology Review*, 6(3), 232–241. https://doi.org/10.1207/S15327957PSPR0603_7
- Brookings. (2023). *The pandemic has had devastating impacts on learning: what will it take to help students catch up?* <https://www.brookings.edu/articles/the-pandemic-has-had-devastating-impacts-on-learning-what-will-it-take-to-help-students-catch-up/>
- Çalışkan, K., & Callon, M. (2009). Economization, part 1: Shifting attention from the economy towards processes of economization. *Economy and Society*, 38(3), 369–398. <https://doi.org/10.1080/03085140903020580>
- Chiapello, E. (2015). Financialisation of Valuation. *Human Studies*, 38, 13–35. <https://doi.org/10.1007/s10746-014-9337-x>
- Cooper, D. (1998). *Governing out of order: Space*. Rivers Oram Press.
- Elfert, M., & Ydesen, C. (2023). *Global governance of education: The historical and contemporary entanglements of UNESCO, the OECD and the World Bank*. Springer.
- Eyal, G. (2019). *The crisis of expertise*. Polity Press.
- Fontdevila, C. (2023). The politics of good enough data. *Developments, Dilemmas and Deadlocks in the Production of Global Learning Metrics, International Journal of Educational Development*, 96. <https://doi.org/10.1016/j.ijedudev.2022.102684>

- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (2010). *The new production of knowledge: The dynamics of science and research in contemporary societies*. SAGE Publications Ltd, <https://doi.org/10.4135/9781446221853>
- Glewwe, P. (2002). Schools and skills in developing countries: Education policies and socioeconomic outcomes. *Journal of Economic Literature*, 40(2), 436–482.
- Grek, S. (2008). From symbols to numbers: The shifting technologies of education governance in Europe. *European Education Research Journal*, 7(2), 208–218.
- Grek, S., Maroy, C., & Verger, A. (2020). *World yearbook of education 2021: Accountability and datafication in the governance of education*. Routledge.
- Grek, S. (2013). Expert moves: International comparative testing and the rise of expertocracy. *Journal of Education Policy*. <https://doi.org/10.1080/02680939.2012.758825>
- Grek, S. (2020). Prophets, saviours and saints: Symbolic governance and the rise of a transnational metrological field. *International Review of Education*. <https://doi.org/10.1007/s11159-020-09844-z>
- Grek, S. (2022 December). The education sustainable development goal and the generative power of failing metrics. *Policy and Society*, 41(4), 445–457. <https://doi.org/10.1093/polsoc/puac020>
- Gümüşay, A. A., & Reinecke, J. (2021). “Researching for Desirable Futures: From Real Utopias to Imagining Alternatives.” *Journal of Management Studies*, 59(1), 236–242.
- Hanushek, E., & Kimko, D. (2000). Schooling, labor force quality and the growth of nations. *American Economic Review*, 95(5), 1184–1208.
- Hillstrom, L. C. (2019). *The #MeToo movement*, ABC-CLIO.
- Jameson, F. (2004) The politics of Utopia. *New Left Review*, 35–54.
- Jasanoff, S. (Ed.). (2004). *States of knowledge: The Co-production of science and the social order*. Routledge.
- Kingsbury, B., Merry, S. E., & Davis, K. E. (2012). Indicators as a technology of global governance. *Law & Society Review*, 46(1), 71–104.
- Kurunmäki, L., Mennicken, A., Miller, P. (2016). Quantifying, economising, and marketising: Democratising the social sphere? *Sociologie du travail*, 58(4), 390–402.
- Lawn, M., & Grek, S. (2012). *Europeanising education: Governing a new policy space*. Symposium.
- Levitas, R. (1990). *The Concept of Utopia: Utopianism and communitarianism*. Syracuse University Press.
- Levitas, R. (2013). *Utopia as method: The imaginary reconstitution of society*. Springer.

- Lövbrand, E. (2011). Co-producing European climate science and policy: A cautionary note on the making of useful knowledge. *Science and Public Policy*, 38(3), 225–236.
- 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>
- Mehrpouya, A., & Samiolo, R. (2016). Performance measurement in global governance: Ranking and the politics of variability. *Accounting, Organisations and Society*, 55, 12–36. <https://doi.org/10.1016/j.aos.2016.09.001>
- Menicken, A., & Espeland, W. (2019). What's new with numbers? *Sociological Approaches to the Study of Quantification, Annual Review of Sociology*, 45(1), 223–245.
- Miller, P., & Power, M. (2013). Accounting, organizing, and economizing: Connecting accounting research and organization theory. *The Academy of Management Annals*, 7(1), 557–605. <https://doi.org/10.1080/19416520.2013.783668>
- Miller, R. (2018). *Transforming the future: Anticipation in the 21st century*. Routledge.
- Montoya, S. (2017). A pragmatic and unified approach to measure learning globally, *UIS Blog*, <https://uis.unesco.org/en/blog/pragmatic-and-unified-approach-measure-learning-globally>. Accessed on 16 July 2023.
- Mundy, K., Green, A., Lingard, B., & Verger, A. (2016). *Handbook of global education policy*. Wiley-Blackwell.
- Muniesa, F., Doganova, L., Ortiz, H., Pina-Stranger, Á., Paterson, F., Bourgoin, A., Ehrenstein, V., Juven, P.-A., Pontille, D., Saraç-Lesavre, B. & Yon, G. (2017). *Capitalization: A cultural guide*. Presses des Mines.
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science: Knowledge and the public in an age of uncertainty*. Polity.
- Nowotny, H., Scott, P., & Gibbons, M. (2003). Introduction: 'Mode 2' Revisited: The new production of knowledge. *Minerva*, 41(3), 179–194.
- Ozga, J., Dahler-Larsen, P., Segerholm, C., & Simola, H. (2011). *Fabricating quality in education: Data and governance in Europe*. Routledge.
- Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.
- Roe, E. (1994). *Narrative policy analysis*. Duke University Press.
- Sillince, J. A., Jarzabkowski, P., & Shaw, D. (2012). Shaping strategic action through the rhetorical construction and exploitation of ambiguity. *Organization Science*, 23(3), 630–650. <https://doi.org/10.1287/orsc.1110.0670IS.N1047-7039>
- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 377–391.

- Star, L. S. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values*, 35(5), 601–617. <https://doi.org/10.1177/0162243910377624>
- Strickland, C. (2022). The fight for equality continues: A new social movement analysis of the black lives matter movement and the 1960's Civil rights movement. *Phylon (1960-)*, 59(1), 71–90. <https://www.jstor.org/stable/27150915>
- Thaler, M. (2019). Bleak Dreams. *Not nightmares: Critical dystopias and the necessity of melancholic hope, Constellations*, 26(4), 607–622.
- Thaler, M. (2022). *No Other Planet: Utopian Visions for a Climate-Changed World*. Cambridge University Press.
- Tichenor, Marlee, Sally E. Merry, Sotiria G., & Justyna, Bandola-Gill. (2022 December). Global public policy in a quantified world: Sustainable development goals as epistemic infrastructures. *Policy and Society*, 41(4), 431–444. <https://doi.org/10.1093/polsoc/puac015>
- 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.
- Wright, E. O. (2010). *Envisioning real Utopias*. Verso.
- Wyborn, C., Datta, A., Montana, J., Ryan, M., Leith, P., Chaffin, B., Miller, C., & van Kerkhoff, L. (2019). Co-Producing Sustainability: Reordering the Governance of Science, Policy, and Practice. *Annual Review of Environment and Resources*, 44(1), 319–346.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



INDEX

A

Accountability, 5, 7, 9, 15, 16
Alignment, 141
Alignment device, 141

B

Belonging, 187
Boundary object, 140
Brokerage, 73, 80, 185, 200
Brokers, 15, 89, 99, 118, 161, 178

C

Collective puzzling, 89, 90, 92, 94,
99, 114
Competition, 153, 154, 157, 158,
163, 178
Consensus, 89–91, 96, 102–104,
106–110, 112, 115–117, 125,
129, 139, 141, 142
Context of application, 185, 186,
188, 192
Controversy, 28, 29

D

Data, 95–97, 99, 104, 105, 107–109,
112, 115–117, 153, 156–162,
165, 167, 169, 170, 174–178
Datafication, 193, 194
Data visualisations, 125, 127,
129–132, 134, 136, 140–145

E

Economics, 61, 65, 69, 75, 78–82
Economisation, 61, 62, 64, 185, 189,
190
Education, 183–191, 193–201
Education for All (EFA), 39–43, 45,
47
Education governance, 2, 6, 16
Epistemic infrastructures, 62–65, 68,
69, 78, 79
European Commission (EC), 25, 30,
38, 49–51
European Education Area (EEA), 3,
183, 184
Europeanisation, 154, 161, 162

Expertise, 2, 4, 6, 9, 10, 15, 16, 184,
185, 188, 189, 191, 197, 200

G

Global challenges, 1, 11

Goal-setting, 90

H

Higher education (HE), 153, 154,
157–160, 163–166, 168–177

I

Ignorance, 66, 79, 81–83

Incheon Declaration, 38, 44–48

Institutionalisation, 27, 28, 33

Interdependence, 38, 49, 51, 52, 54

Interdependency, 25, 30, 36, 52, 90

Interdisciplinarity, 61

International Adult Literacy Survey
(IALS), 31, 33

International comparative assessments,
31

International Organisations (IOs),
1–3, 12, 15

K

Knowledge, 1–16

Knowledge controversies, 25, 28

L

Learning, 63–65, 69, 71–74, 77, 78

Learning assessments, 153, 154,
156–158, 177

M

Market, 153, 154, 156–161, 165,
169–173, 175–178

Market of measures, 185, 194, 195

Materiality, 67, 68, 71, 79

Measurement, 1, 3–5, 8, 14, 15, 17,
153, 154, 157, 158, 160, 161,
165, 170, 171, 173, 176–178

Metrology, 25

Mode 1/Mode 2, 6, 11–16

Monitoring, 2–5, 7

Mono-disciplinarity, 62, 66, 75, 78,
82, 185, 188–190, 200

Muscat Agreement, 38, 40–44, 46,
47, 53, 54

N

Narrative-making, 38, 41

Non-knowledge, 62, 66, 75, 77, 78,
81–83

Numbers, 2–6

O

OECD, 25, 30–38, 43, 45, 49–51,
62, 64, 69–71, 73, 80

OECD country reviews, 90, 95

P

Performance management, 187, 194

Political accountability, 89, 107, 109,
116

Political work, 28, 38, 52

Problematization, 3

Programme for International

Assessment of Adult

Competencies (PIAAC), 32, 33,
36

Programme for International Student
Assessment (PISA), 24, 31–34,
38

Provisional numbers, 192, 200

Q

Qualification, 126, 128, 147

Qualification, 125, 127–129, 146, 147
 Quality, 159, 160, 162, 163, 166, 167, 169, 170, 174, 176
 Quality assurance (QA), 154, 158–160, 163, 165, 166, 170–173, 176, 177
 Quantification, 1–3, 5, 10, 11, 15–17, 125, 127–129, 137, 140, 143, 145–147, 183, 184, 187, 190, 192, 194, 196, 198–201

R

Reflexivity, 124, 125, 129–131, 136, 137, 139–147, 185, 192–194, 197, 200
 Re-politicisation, 5, 114, 118

S

Socialisation, 89–92, 94, 96, 98, 99, 113–115
 Standards, 160, 162, 163, 166, 167, 169, 172
 Storytelling, 124, 125, 128, 130, 131, 136, 140, 143, 144, 147

Sustainable Development Goal 4 (SDG4), 63–67, 72–75, 78–81
 Sustainable Development Goals (SDGs), 3, 4, 24, 38, 40, 42–47, 54, 183, 191, 197
 Sweden, 91, 94–99, 114

T

Technicisation, 113
 Technocratic accountability, 103
 Transdisciplinarity, 15, 185, 188
 Transnational governance, 2, 7, 15

U

United Nations Educational, Scientific and Cultural Organization (UNESCO), 62, 63, 71–75
 Universality, 24, 39, 41, 53, 54, 187, 188
 UN participatory turn, 90, 99, 100, 113, 115
 Utopia, 185, 196, 198, 199, 201

W

World Bank, 62, 70–73, 80